

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
5 April 2001 (05.04.2001)

PCT

(10) International Publication Number
WO 01/23416 A2

(51) International Patent Classification⁷: C07K 14/00

(21) International Application Number: PCT/EP00/09495

(22) International Filing Date:
26 September 2000 (26.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9923156.5 30 September 1999 (30.09.1999) GB

(71) Applicant (for all designated States except US):
SMITHKLINE BEECHAM BIOLOGICALS S.A.
[BE/BE]; Rue de l'Institut 89, B-1330 Rixensart (BE).

(72) Inventor; and

(75) Inventor/Applicant (for US only): THONNARD, Joelle
[BE/BE]; SmithKline Beecham Biologicals s.a., Rue de
l'Institut 89, B-1330 Rixensart (BE).

(74) Agent: PRIVETT, Kathryn, Louise; SmithKline
Beecham, Two New Horizons Court, Brentford, Middlesex
TW8 9EP (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— Without international search report and to be republished
upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: NOVEL COMPOUNDS

(57) Abstract: The invention provides BASB132 polypeptides and polynucleotides encoding BASB132 polypeptides and methods for producing such polypeptides by recombinant techniques. Also provided are diagnostic, prophylactic and therapeutic uses.

WO 01/23416 A2

Novel Compounds

FIELD OF THE INVENTION

This invention relates to polynucleotides, (herein referred to as “BASB132 polynucleotide(s)”), polypeptides encoded by them (referred to herein as “BASB132” or “BASB132 polypeptide(s)”), recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including vaccines against bacterial infections. In a further aspect, the invention relates to diagnostic assays for detecting infection of certain pathogens.

BACKGROUND OF THE INVENTION

Moraxella catarrhalis (also named *Branhamella catarrhalis*) is a Gram negative bacteria frequently isolated from the human upper respiratory tract. It is responsible for several pathologies the main ones being otitis media in infants and children, and pneumonia in elderlies. It is also responsible of sinusitis, nosocomial infections and less frequently of invasive diseases.

Otitis media is an important childhood disease both by the number of cases and its potential sequelae. More than 3.5 millions cases are recorded every year in the United States, and it is estimated that 80 % of the children have experienced at least one episode of otitis before reaching the age of 3 (Klein, JO (1994) Clin.Inf.Dis 19:823). Left untreated, or becoming chronic, this disease may lead to hearing losses that could be temporary (in the case of fluid accumulation in the middle ear) or permanent (if the auditive nerve is damaged). In infants, such hearing losses may be responsible for a delayed speech learning.

Three bacterial species are primarily isolated from the middle ear of children with otitis media: *Streptococcus pneumoniae*, non typeable *Haemophilus influenza* (NTHi) and *M. catarrhalis*. They are present in 60 to 90 % of the cases. A review of recent studies shows that *S. pneumoniae* and NTHi represent both about 30 %, and *M. catarrhalis* about 15 % of

the otitis media cases (Murphy, TF (1996) Microbiol.Rev. 60:267). Other bacteria could be isolated from the middle ear (*H. influenza* type B, *S. pyogenes* etc) but at a much lower frequency (2 % of the cases or less).

Epidemiological data indicate that, for the pathogens found in the middle ear, the colonization of the upper respiratory tract is an absolute prerequisite for the development of an otitis; other are however also required to lead to the disease (Dickinson, DP et al. (1988) J. Infect.Dis. 158:205, Faden, HL et al. (1991) Ann.Otorhinol.Laryngol. 100:612). These are important to trigger the migration of the bacteria into the middle ear via the Eustachian tubes, followed by the initiation of an inflammatory process. These factors are unknown todate. It has been postulated that a transient anomaly of the immune system following a viral infection, for example, could cause an inability to control the colonization of the respiratory tract (Faden, HL et al (1994) J. Infect.Dis. 169:1312). An alternative explanation is that the exposure to environmental factors allow a more important colonization of some children, who subsequently become susceptible to the development of otitis media because of the sustained presence of middle ear pathogens (Murphy, TF (1996) Microbiol.Rev. 60:267).

The immune response to *M. catarrhalis* is poorly characterized. The analysis of strains isolated sequentially from the nasopharynx of babies followed from 0 to 2 years of age, indicates that they get and eliminate frequently new strains. This indicates that an efficacious immune response against this bacteria is mounted by the colonized children (Faden, HL et al (1994) J. Infect.Dis. 169:1312).

In most adults tested, bactericidal antibodies have been identified (Chapman, AJ et al. (1985) J. Infect.Dis. 151:878). Strains of *M. catarrhalis* present variations in their capacity to resist serum bactericidal activity: in general, isolates from diseased individuals are more resistant than those who are simply colonized (Hol, C et al. (1993) Lancet 341:1281, Jordan, KL et al. (1990) Am.J.Med. 88 (suppl. 5A):28S). Serum resistance could therefore be

considered as a virulence factor of the bacteria. An opsonizing activity has been observed in the sera of children recovering from otitis media.

The antigens targetted by these different immune responses in humans have not been identified, with the exception of OMP B1, a 84 kDa protein which expression is regulated by iron, and that is recognized by the sera of patients with pneumonia (Sethi, S, et al. (1995) Infect.Immun. 63:1516) , and of UspA1 and UspA2 (Chen D. et al.(1999), Infect.Immun. 67:1310).

A few other membrane proteins present on the surface of *M. catarrhalis* have been characterized using biochemical method, or for their potential implication in the induction of a protective immunity (for review, see Murphy, TF (1996) Microbiol.Rev. 60:267). In a mouse pneumonia model, the presence of antibodies raised against some of them (UspA, CopB) favors a faster clearance of the pulmonary infection. Another polypeptide (OMP CD) is highly conserved among *M. catarrhalis* strains, and presents homologies with a porin of *Pseudomonas aeruginosa*, which has been demonstrated efficacious against this bacterium in animal models.

The frequency of *Moraxella catarrhalis* infections has risen dramatically in the past few decades. This has been attributed to the emergence of multiply antibiotic resistant strains and an increasing population of people with weakened immune systems. It is no longer uncommon to isolate *Moraxella catarrhalis* strains that are resistant to some or all of the standard antibiotics. This phenomenon has created an unmet medical need and demand for new anti-microbial agents, vaccines, drug screening methods, and diagnostic tests for this organism.

SUMMARY OF THE INVENTION

The present invention relates to BASB132, in particular BASB132 polypeptides and BASB132 polynucleotides, recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including prevention and treatment of microbial diseases, amongst others. In a further aspect, the invention relates to diagnostic assays for detecting diseases associated with microbial infections and conditions associated with such infections, such as assays for detecting expression or activity of BASB132 polynucleotides or polypeptides.

Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

DESCRIPTION OF THE INVENTION

The invention relates to BASB132 polypeptides and polynucleotides as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of BASB132 of *Moraxella catarrhalis*, which is related by amino acid sequence homology to *Escherichia coli* hypothetical outer membrane protein YtfN. The invention relates especially to BASB132 having the nucleotide and amino acid sequences set out in SEQ ID NO:1 or 3 and SEQ ID NO:2 or 4 respectively. It is understood that sequences recited in the Sequence Listing below as "DNA" represent an exemplification of one embodiment of the invention, since those of ordinary skill will recognize that such sequences can be usefully employed in polynucleotides in general, including ribopolynucleotides.

Polypeptides

In one aspect of the invention there are provided polypeptides of *Moraxella catarrhalis* referred to herein as "BASB132" and "BASB132 polypeptides" as well as biologically,

diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

The present invention further provides for:

- (a) an isolated polypeptide which comprises an amino acid sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, most preferably at least 97-99% or exact identity, to that of SEQ ID NO:2 or 4;
- (b) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1 or 3 over the entire length of SEQ ID NO:1 or 3 respectively; or
- (c) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or exact identity, to the amino acid sequence of SEQ ID NO:2 or 4.

The BASB132 polypeptides provided in SEQ ID NO:2 or 4 are the BASB132 polypeptides from *Moraxella catarrhalis* strain Mc2931 (ATCC 43617):

The invention also provides an immunogenic fragment of a BASB132 polypeptide, that is, a contiguous portion of the BASB132 polypeptide which has the same or substantially the same immunogenic activity as the polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 4; That is to say, the fragment (if necessary when coupled to a carrier) is capable of raising an immune response which recognises the BASB132 polypeptide. Such an immunogenic fragment may include, for example, the BASB132 polypeptide lacking an N-terminal leader sequence, and/or a transmembrane domain and/or a C-terminal anchor domain. In a preferred aspect the immunogenic fragment of BASB132 according to the invention comprises substantially all of the extracellular domain of a polypeptide which has at least 85% identity, preferably at least 90% identity, more

preferably at least 95% identity, most preferably at least 97-99% identity, to that of SEQ ID NO:2 or 4 over the entire length of SEQ ID NO:2 or 4 respectively.

A fragment is a polypeptide having an amino acid sequence that is entirely the same as part but not all of any amino acid sequence of any polypeptide of the invention. As with BASB132 polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region in a single larger polypeptide.

Preferred fragments include, for example, truncation polypeptides having a portion of an amino acid sequence of SEQ ID NO:2 or 4 or of variants thereof, such as a continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence. Degradation forms of the polypeptides of the invention produced by or in a host cell, are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

Further preferred fragments include an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids from the amino acid sequence of SEQ ID NO:2 or 4, or an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids truncated or deleted from the amino acid sequence of SEQ ID NO:2 or 4.

Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these fragments

may be employed as intermediates for producing the full-length polypeptides of the invention.

Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added in any combination.

The polypeptides, or immunogenic fragments, of the invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.

In one aspect, the invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa.

Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

The proteins may be chemically conjugated, or expressed as recombinant fusion proteins allowing increased levels to be produced in an expression system as compared to non-fused protein. The fusion partner may assist in providing T helper epitopes (immunological fusion partner), preferably T helper epitopes recognised by humans, or assist in expressing the protein (expression enhancer) at higher yields than the native recombinant protein. Preferably the fusion partner will be both an immunological fusion partner and expression enhancing partner.

Fusion partners include protein D from *Haemophilus influenzae* and the non-structural protein from influenza virus, NS1 (hemagglutinin). Another fusion partner is the protein known as LytA. Preferably the C terminal portion of the molecule is used. LytA is derived from *Streptococcus pneumoniae* which synthesize an N-acetyl-L-alanine amidase, amidase LytA, (coded by the lytA gene {Gene, 43 (1986) page 265-272}) an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LytA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of E.coli C-LytA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LytA fragment at its amino terminus has been described {Biotechnology: 10, (1992) page 795-798}. It is possible to use the repeat portion of the LytA molecule found in the C terminal end starting at residue 178, for example residues 188 - 305.

The present invention also includes variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions, whereby a residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

It is most preferred that a polypeptide of the invention is derived from *Moraxella catarrhalis*, however, it may preferably be obtained from other organisms of the same taxonomic genus. A polypeptide of the invention may also be obtained, for example, from organisms of the same taxonomic family or order.

Polynucleotides

It is an object of the invention to provide polynucleotides that encode BASB132 polypeptides, particularly polynucleotides that encode the polypeptide herein designated BASB132.

In a particularly preferred embodiment of the invention the polynucleotide comprises a region encoding BASB132 polypeptides comprising a sequence set out in SEQ ID NO:1 or 3 which includes a full length gene, or a variant thereof.

The BASB132 polynucleotides provided in SEQ ID NO:1 or 3 are the BASB132 polynucleotides from *Moraxella catarrhalis* strain Mc2931 (ATCC 43617).

As a further aspect of the invention there are provided isolated nucleic acid molecules encoding and/or expressing BASB132 polypeptides and polynucleotides, particularly *Moraxella catarrhalis* BASB132 polypeptides and polynucleotides, including, for example, unprocessed RNAs, ribozyme RNAs, mRNAs, cDNAs, genomic DNAs, B- and Z-DNAs. Further embodiments of the invention include biologically,

diagnostically, prophylactically, clinically or therapeutically useful polynucleotides and polypeptides, and variants thereof, and compositions comprising the same.

Another aspect of the invention relates to isolated polynucleotides, including at least one full length gene, that encodes a BASB132 polypeptide having a deduced amino acid sequence of SEQ ID NO:2 or 4 and polynucleotides closely related thereto and variants thereof.

In another particularly preferred embodiment of the invention there is a BASB132 polypeptide from *Moraxella catarrhalis* comprising or consisting of an amino acid sequence of SEQ ID NO:2 or 4 or a variant thereof.

Using the information provided herein, such as a polynucleotide sequence set out in SEQ ID NO:1 or 3, a polynucleotide of the invention encoding BASB132 polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and sequencing chromosomal DNA fragments from bacteria using *Moraxella catarrhalis* Catlin cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide sequence of the invention, such as a polynucleotide sequence given in SEQ ID NO:1 or 3, typically a library of clones of chromosomal DNA of *Moraxella catarrhalis* Catlin in *E.coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent hybridization conditions. By sequencing the individual clones thus identified by hybridization with sequencing primers designed from the original polypeptide or polynucleotide sequence it is then possible to extend the polynucleotide sequence in both directions to determine a full length gene sequence. Conveniently, such sequencing is performed, for example, using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). (see in particular Screening By

Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Direct genomic DNA sequencing may also be performed to obtain a full length gene sequence. Illustrative of the invention, each polynucleotide set out in SEQ ID NO:1 or 3 was discovered in a DNA library derived from *Moraxella catarrhalis*.

Moreover, each DNA sequence set out in SEQ ID NO:1 or 3 contains an open reading frame encoding a protein having about the number of amino acid residues set forth in SEQ ID NO:2 or 4 with a deduced molecular weight that can be calculated using amino acid residue molecular weight values well known to those skilled in the art.

The polynucleotide of SEQ ID NO:1, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 5017 of SEQ ID NO:1, encodes the polypeptide of SEQ ID NO:2.

The polynucleotide of SEQ ID NO:3, between the start codon at nucleotide number 2042 and the stop codon which begins at nucleotide number 5014 of SEQ ID NO:3, encodes the polypeptide of SEQ ID NO:4.

In a further aspect, the present invention provides for an isolated polynucleotide comprising or consisting of:

- (a) a polynucleotide sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1 or 3 over the entire length of SEQ ID NO:1 or 3 respectively; or
- (b) a polynucleotide sequence encoding a polypeptide which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or 100% exact, to the amino acid sequence of SEQ ID NO:2 or 4, over the entire length of SEQ ID NO:2 or 4 respectively.

A polynucleotide encoding a polypeptide of the present invention, including homologs and orthologs from species other than *Moraxella catarrhalis*, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions (for example, using a temperature in the range of 45 – 65°C and an SDS concentration from 0.1 – 1%) with a labeled or detectable probe consisting of or comprising the sequence of SEQ ID NO:1 or 3 or a fragment thereof; and isolating a full-length gene and/or genomic clones containing said polynucleotide sequence.

The invention provides a polynucleotide sequence identical over its entire length to a coding sequence (open reading frame) in SEQ ID NO:1 or 3. Also provided by the invention is a coding sequence for a mature polypeptide or a fragment thereof, by itself as well as a coding sequence for a mature polypeptide or a fragment in reading frame with another coding sequence, such as a sequence encoding a leader or secretory sequence, a pre-, or pro- or prepro-protein sequence. The polynucleotide of the invention may also contain at least one non-coding sequence, including for example, but not limited to at least one non-coding 5' and 3' sequence, such as the transcribed but non-translated sequences, termination signals (such as rho-dependent and rho-independent termination signals), ribosome binding sites, Kozak sequences, sequences that stabilize mRNA, introns, and polyadenylation signals. The polynucleotide sequence may also comprise additional coding sequence encoding additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA* 86: 821-824 (1989), or an HA peptide tag (Wilson *et al.*, *Cell* 37: 767 (1984), both of which may be useful in purifying polypeptide sequence fused to them. Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

The nucleotide sequence encoding BASB132 polypeptide of SEQ ID NO:2 or 4 may be identical to the polypeptide encoding sequence contained in nucleotides 1 to 5016 of SEQ ID NO:1 or the polypeptide encoding sequence contained in nucleotides 1 to 5016 of SEQ ID NO:3 respectively. Alternatively it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:2 or 4.

The term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides that include a sequence encoding a polypeptide of the invention, particularly a bacterial polypeptide and more particularly a polypeptide of the *Moraxella catarrhalis* BASB132 having an amino acid sequence set out in SEQ ID NO:2 or 4. The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, polynucleotides interrupted by integrated phage, an integrated insertion sequence, an integrated vector sequence, an integrated transposon sequence, or due to RNA editing or genomic DNA reorganization) together with additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the polynucleotides described herein that encode variants of a polypeptide having a deduced amino acid sequence of SEQ ID NO:2 or 4. Fragments of polynucleotides of the invention may be used, for example, to synthesize full-length polynucleotides of the invention.

Further particularly preferred embodiments are polynucleotides encoding BASB132 variants, that have the amino acid sequence of BASB132 polypeptide of SEQ ID NO:2 or 4 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, modified, deleted and/or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, that do not alter the properties and activities of BASB132 polypeptide.

Further preferred embodiments of the invention are polynucleotides that are at least 85% identical over their entire length to a polynucleotide encoding BASB132 polypeptide having an amino acid sequence set out in SEQ ID NO:2 or 4, and polynucleotides that are complementary to such polynucleotides. Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 90% identical over its entire length to a polynucleotide encoding BASB132 polypeptide and polynucleotides complementary thereto. In this regard, polynucleotides at least 95% identical over their entire length to the same are particularly preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

Preferred embodiments are polynucleotides encoding polypeptides that retain substantially the same biological function or activity as the mature polypeptide encoded by a DNA of SEQ ID NO:1 or 3.

In accordance with certain preferred embodiments of this invention there are provided polynucleotides that hybridize, particularly under stringent conditions, to BASB132 polynucleotide sequences, such as those polynucleotides in SEQ ID NO:1 or 3.

The invention further relates to polynucleotides that hybridize to the polynucleotide sequences provided herein. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the polynucleotides described herein. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization occurring only if there is at least 95% and preferably at least 97% identity between the sequences. A specific example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml of denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at about 65°C.

Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein. Solution hybridization may also be used with the polynucleotide sequences provided by the invention.

The invention also provides a polynucleotide consisting of or comprising a polynucleotide sequence obtained by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in SEQ ID NO:1 or 3 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence set forth in SEQ ID NO:1 or 3 or a fragment thereof; and isolating said polynucleotide sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers fully described elsewhere herein.

As discussed elsewhere herein regarding polynucleotide assays of the invention, for instance, the polynucleotides of the invention, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding BASB132 and to isolate cDNA and genomic clones of other genes that have a high identity, particularly high sequence identity, to the BASB132 gene. Such probes generally will comprise at least 15 nucleotide residues or base pairs. Preferably, such probes will have at least 30 nucleotide residues or base pairs and may have at least 50 nucleotide residues or base pairs. Particularly preferred probes will have at least 20 nucleotide residues or base pairs and will have less than 30 nucleotide residues or base pairs.

A coding region of a BASB132 gene may be isolated by screening using a DNA sequence provided in SEQ ID NO:1 or 3 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

There are several methods available and well known to those skilled in the art to obtain full-length DNAs, or extend short DNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman, *et al.*, *PNAS USA* 85: 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the DNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using "nested" primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the selected gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length DNA constructed either by joining the product directly to the existing DNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for diseases, particularly human diseases, as further discussed herein relating to polynucleotide assays.

The polynucleotides of the invention that are oligonucleotides derived from a sequence of SEQ ID NOS:1 or 3 may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The invention also provides polynucleotides that encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

For each and every polynucleotide of the invention there is provided a polynucleotide complementary to it. It is preferred that these complementary polynucleotides are fully complementary to each polynucleotide with which they are complementary.

A precursor protein, having a mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In addition to the standard A, G, C, T/U representations for nucleotides, the term "N" may also be used in describing certain polynucleotides of the invention. "N" means that any of the four DNA or RNA nucleotides may appear at such a designated position in the DNA or RNA sequence, except it is preferred that N is not a nucleic acid that when taken in combination with adjacent nucleotide positions, when read in the correct reading frame, would have the effect of generating a premature termination codon in such reading frame.

In sum, a polynucleotide of the invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or

more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

In accordance with an aspect of the invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff *et al.*, *Hum Mol Genet* (1992) 1: 363, Manthorpe *et al.*, *Hum. Gene Ther.* (1983) 4: 419), delivery of DNA complexed with specific protein carriers (Wu *et al.*, *J Biol Chem.* (1989) 264: 16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS USA*, (1986) 83: 9551), encapsulation of DNA in various forms of liposomes (Kaneda *et al.*, *Science* (1989) 243: 375), particle bombardment (Tang *et al.*, *Nature* (1992) 356:152, Eisenbraun *et al.*, *DNA Cell Biol* (1993) 12: 791) and *in vivo* infection using cloned retroviral vectors (Seeger *et al.*, *PNAS USA* (1984) 81: 5849).

Vectors, Host Cells, Expression Systems

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

Recombinant polypeptides of the present invention may be prepared by processes well known in those skilled in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems that comprise a polynucleotide or polynucleotides of the present

invention, to host cells which are genetically engineered with such expression systems, and to the production of polypeptides of the invention by recombinant techniques.

For recombinant production of the polypeptides of the invention, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis, *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook, *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as cells of streptococci, staphylococci, enterococci, *E. coli*, streptomyces, cyanobacteria, *Bacillus subtilis*, *Neisseria meningitidis* and *Moraxella catarrhalis*; fungal cells, such as cells of a yeast, *Kluveromyces*, *Saccharomyces*, a basidiomycete, *Candida albicans* and *Aspergillus*; insect cells such as cells of *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293, CV-1 and Bowes melanoma cells; and plant cells, such as cells of a gymnosperm or angiosperm.

A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal-, episomal- and virus-derived vectors, for example, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses, picornaviruses, retroviruses, and alphaviruses and vectors derived from combinations thereof, such as those derived from

plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

In recombinant expression systems in eukaryotes, for secretion of a translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, ion metal affinity chromatography (IMAC) is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

The expression system may also be a recombinant live microorganism, such as a virus or bacterium. The gene of interest can be inserted into the genome of a live recombinant virus or bacterium. Inoculation and *in vivo* infection with this live vector will lead to *in vivo* expression of the antigen and induction of immune responses. Viruses and bacteria used for this purpose are for instance: poxviruses (e.g; vaccinia, fowlpox, canarypox), alphaviruses (Sindbis virus, Semliki Forest Virus, Venezuelan Equine Encephalitis

Virus), adenoviruses, adeno-associated virus, picornaviruses (poliovirus, rhinovirus), herpesviruses (varicella zoster virus, etc), Listeria, Salmonella, Shigella, BCG. These viruses and bacteria can be virulent, or attenuated in various ways in order to obtain live vaccines. Such live vaccines also form part of the invention.

Diagnostic, Prognostic, Serotyping and Mutation Assays

This invention is also related to the use of BASB132 polynucleotides and polypeptides of the invention for use as diagnostic reagents. Detection of BASB132 polynucleotides and/or polypeptides in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of disease, staging of disease or response of an infectious organism to drugs. Eukaryotes, particularly mammals, and especially humans, particularly those infected or suspected to be infected with an organism comprising the BASB132 gene or protein, may be detected at the nucleic acid or amino acid level by a variety of well known techniques as well as by methods provided herein.

Polypeptides and polynucleotides for prognosis, diagnosis or other analysis may be obtained from a putatively infected and/or infected individual's bodily materials. Polynucleotides from any of these sources, particularly DNA or RNA, may be used directly for detection or may be amplified enzymatically by using PCR or any other amplification technique prior to analysis. RNA, particularly mRNA, cDNA and genomic DNA may also be used in the same ways. Using amplification, characterization of the species and strain of infectious or resident organism present in an individual, may be made by an analysis of the genotype of a selected polynucleotide of the organism. Deletions and insertions can be detected by a change in size of the amplified product in comparison to a genotype of a reference sequence selected from a related organism, preferably a different species of the same genus or a different strain of the same species. Point mutations can be identified by hybridizing amplified DNA to labeled BASB132 polynucleotide sequences. Perfectly or significantly matched sequences can be distinguished from imperfectly or more significantly mismatched duplexes by DNase or RNase digestion, for DNA or RNA respectively, or by detecting

differences in melting temperatures or renaturation kinetics. Polynucleotide sequence differences may also be detected by alterations in the electrophoretic mobility of polynucleotide fragments in gels as compared to a reference sequence. This may be carried out with or without denaturing agents. Polynucleotide differences may also be detected by direct DNA or RNA sequencing. See, for example, Myers *et al.*, *Science*, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase, V1 and S1 protection assay or a chemical cleavage method. See, for example, Cotton *et al.*, *Proc. Natl. Acad. Sci., USA*, 85: 4397-4401 (1985).

In another embodiment, an array of oligonucleotides probes comprising BASB132 nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of, for example, genetic mutations, serotype, taxonomic classification or identification. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see, for example, Chee *et al.*, *Science*, 274: 610 (1996)).

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO:1 or 3, or a fragment thereof;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2 or 4 or a fragment thereof; or
- (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2 or 4.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a Disease, among others.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of a polynucleotide of the invention, preferably SEQ ID NO:1 or 3, which is associated with a disease or pathogenicity will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, a prognosis of a course of disease, a determination of a stage of disease, or a susceptibility to a disease, which results from under-expression, over-expression or altered expression of the polynucleotide. Organisms, particularly infectious organisms, carrying mutations in such polynucleotide may be detected at the polynucleotide level by a variety of techniques, such as those described elsewhere herein.

Cells from an organism carrying mutations or polymorphisms (allelic variations) in a polynucleotide and/or polypeptide of the invention may also be detected at the polynucleotide or polypeptide level by a variety of techniques, to allow for serotyping, for example. For example, RT-PCR can be used to detect mutations in the RNA. It is particularly preferred to use RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA, cDNA or genomic DNA may also be used for the same purpose, PCR. As an example, PCR primers complementary to a polynucleotide encoding BASB132 polypeptide can be used to identify and analyze mutations.

The invention further provides primers with 1, 2, 3 or 4 nucleotides removed from the 5' and/or the 3' end. These primers may be used for, among other things, amplifying BASB132 DNA and/or RNA isolated from a sample derived from an individual, such as a bodily material. The primers may be used to amplify a polynucleotide isolated from an infected individual, such that the polynucleotide may then be subject to various techniques for elucidation of the polynucleotide sequence. In this way, mutations in the polynucleotide sequence may be detected and used to diagnose and/or prognose the infection or its stage or course, or to serotype and/or classify the infectious agent.

The invention further provides a process for diagnosing, disease, preferably bacterial infections, more preferably infections caused by *Moraxella catarrhalis*, comprising determining from a sample derived from an individual, such as a bodily material, an increased level of expression of polynucleotide having a sequence of SEQ ID NO:1 or 3. Increased or decreased expression of a BASB132 polynucleotide can be measured using any one of the methods well known in the art for the quantitation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting, spectrometry and other hybridization methods.

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of BASB132 polypeptide compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a BASB132 polypeptide, in a sample derived from a host, such as a bodily material, are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis, antibody sandwich assays, antibody detection and ELISA assays.

The polynucleotides of the invention may be used as components of polynucleotide arrays, preferably high density arrays or grids. These high density arrays are particularly useful for diagnostic and prognostic purposes. For example, a set of spots each comprising a different gene, and further comprising a polynucleotide or polynucleotides of the invention, may be used for probing, such as using hybridization or nucleic acid amplification, using a probes obtained or derived from a bodily sample, to determine the presence of a particular polynucleotide sequence or related sequence in an individual. Such a presence may indicate the presence of a pathogen, particularly *Moraxella catarrhalis*, and may be useful in diagnosing and/or prognosing disease or a course of disease. A grid comprising a number of variants of the polynucleotide sequence of SEQ ID NO:1 or 3 are preferred. Also preferred is a comprising a number

of variants of a polynucleotide sequence encoding the polypeptide sequence of SEQ ID NO:2 or 4.

Antibodies

The polypeptides and polynucleotides of the invention or variants thereof, or cells expressing the same can be used as immunogens to produce antibodies immunospecific for such polypeptides or polynucleotides respectively. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

In certain preferred embodiments of the invention there are provided antibodies against BASB132 polypeptides or polynucleotides.

Antibodies generated against the polypeptides or polynucleotides of the invention can be obtained by administering the polypeptides and/or polynucleotides of the invention, or epitope-bearing fragments of either or both, analogues of either or both, or cells expressing either or both, to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides or polynucleotides of this invention. Also, transgenic mice, or other organisms or animals, such as other mammals, may be used to express humanized antibodies immunospecific to the polypeptides or polynucleotides of the invention.

Alternatively, phage display technology may be utilized to select antibody genes with binding activities towards a polypeptide of the invention either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-BASB132 or from naive libraries (McCafferty, *et al.*, (1990), *Nature* 348, 552-554; Marks, *et al.*, (1992) *Biotechnology* 10, 779-783). The affinity of these antibodies can also be improved by, for example, chain shuffling (Clackson *et al.*, (1991) *Nature* 352: 628).

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides or polynucleotides of the invention to purify the polypeptides or polynucleotides by, for example, affinity chromatography.

Thus, among others, antibodies against BASB132-polypeptide or BASB132-polynucleotide may be employed to treat infections, particularly bacterial infections.

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants form a particular aspect of this invention.

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized," where the complementarity determining region or regions of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones *et al.* (1986), *Nature* 321, 522-525 or Tempest *et al.*, (1991) *Biotechnology* 9, 266-273.

Antagonists and Agonists - Assays and Molecules

Polypeptides and polynucleotides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural

substrates and ligands or may be structural or functional mimetics. See, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology* 1(2): Chapter 5 (1991).

The screening methods may simply measure the binding of a candidate compound to the polypeptide or polynucleotide, or to cells or membranes bearing the polypeptide or polynucleotide, or a fusion protein of the polypeptide by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide or polynucleotide, using detection systems appropriate to the cells comprising the polypeptide or polynucleotide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Constitutively active polypeptide and/or constitutively expressed polypeptides and polynucleotides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate compound results in inhibition of activation of the polypeptide or polynucleotide, as the case may be. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide or polynucleotide of the present invention, to form a mixture, measuring BASB132 polypeptide and/or polynucleotide activity in the mixture, and comparing the BASB132 polypeptide and/or polynucleotide activity of the mixture to a standard. Fusion proteins, such as those made from Fc portion and BASB132 polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists of the polypeptide of the present invention, as well as of phylogenetically and and/or functionally related polypeptides (see D. Bennett *et al.*, *J Mol Recognition*, 8:52-58 (1995); and K. Johanson *et al.*, *J Biol Chem*, 270(16):9459-9471 (1995)).

The polynucleotides, polypeptides and antibodies that bind to and/or interact with a polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and/or polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents which may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of BASB132 polypeptides or polynucleotides, particularly those compounds that are bacteriostatic and/or bactericidal. The method of screening may involve high-throughput techniques. For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising BASB132 polypeptide and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be a BASB132 agonist or antagonist. The ability of the candidate molecule to agonize or antagonize the BASB132 polypeptide is reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Molecules that bind gratuitously, *i.e.*, without inducing the effects of BASB132 polypeptide are most likely to be good antagonists. Molecules that bind well and, as the case may be, increase the rate of product production from substrate, increase signal transduction, or increase chemical channel activity are agonists. Detection of the rate or level of, as the case may be, production of product from substrate, signal transduction, or chemical channel activity may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric, labeled substrate converted into product, a reporter gene that is responsive to changes in BASB132 polynucleotide or polypeptide activity, and binding assays known in the art.

Another example of an assay for BASB132 agonists is a competitive assay that combines BASB132 and a potential agonist with BASB132-binding molecules, recombinant BASB132 binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. BASB132 can be labeled, such as by radioactivity or a colorimetric compound, such that the number of BASB132 molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include, among others, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention and thereby inhibit or extinguish its activity or expression. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing BASB132-induced activities, thereby preventing the action or expression of BASB132 polypeptides and/or polynucleotides by excluding BASB132 polypeptides and/or polynucleotides from binding.

Potential antagonists include a small molecule that binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem.* 56: 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*, CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred potential antagonists include compounds related to and variants of BASB132.

In a further aspect, the present invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of

various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

Each of the polynucleotide sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the polynucleotide sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide, agonist or antagonist of the invention to interfere with the initial physical interaction between a pathogen or pathogens and a eukaryotic, preferably mammalian, host responsible for sequelae of infection. In particular, the molecules of the invention may be used: in the prevention of adhesion of bacteria, in particular gram positive and/or gram negative bacteria, to eukaryotic, preferably mammalian, extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; to block bacterial adhesion between eukaryotic, preferably mammalian, extracellular matrix proteins and bacterial BASB132 proteins that mediate tissue damage and/or; to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

In accordance with yet another aspect of the invention, there are provided BASB132 agonists and antagonists, preferably bacteristatic or bactericidal agonists and antagonists.

The antagonists and agonists of the invention may be employed, for instance, to prevent, inhibit and/or treat diseases.

In a further aspect, the present invention relates to mimotopes of the polypeptide of the invention. A mimotope is a peptide sequence, sufficiently similar to the native peptide (sequentially or structurally), which is capable of being recognised by antibodies which recognise the native peptide; or is capable of raising antibodies which recognise the native peptide when coupled to a suitable carrier.

Peptide mimotopes may be designed for a particular purpose by addition, deletion or substitution of elected amino acids. Thus, the peptides may be modified for the purposes of ease of conjugation to a protein carrier. For example, it may be desirable for some chemical conjugation methods to include a terminal cysteine. In addition it may be desirable for peptides conjugated to a protein carrier to include a hydrophobic terminus distal from the conjugated terminus of the peptide, such that the free unconjugated end of the peptide remains associated with the surface of the carrier protein. Thereby presenting the peptide in a conformation which most closely resembles that of the peptide as found in the context of the whole native molecule. For example, the peptides may be altered to have an N-terminal cysteine and a C-terminal hydrophobic amidated tail. Alternatively, the addition or substitution of a D-stereoisomer form of one or more of the amino acids may be performed to create a beneficial derivative, for example to enhance stability of the peptide.

Alternatively, peptide mimotopes may be identified using antibodies which are capable themselves of binding to the polypeptides of the present invention using techniques such as phage display technology (EP 0 552 267 B1). This technique, generates a large number

of peptide sequences which mimic the structure of the native peptides and are, therefore, capable of binding to anti-native peptide antibodies, but may not necessarily themselves share significant sequence homology to the native polypeptide.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal, preferably humans, which comprises inoculating the individual with BASB132 polynucleotide and/or polypeptide, or a fragment or variant thereof, adequate to produce antibody and/ or T cell immune response to protect said individual from infection, particularly bacterial infection and most particularly *Moraxella catarrhalis* infection. Also provided are methods whereby such immunological response slows bacterial replication. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector, sequence or ribozyme to direct expression of BASB132 polynucleotide and/or polypeptide, or a fragment or a variant thereof, for expressing BASB132 polynucleotide and/or polypeptide, or a fragment or a variant thereof *in vivo* in order to induce an immunological response, such as, to produce antibody and/ or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual, preferably a human, from disease, whether that disease is already established within the individual or not. One example of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a ribozyme, a modified nucleic acid, a DNA/RNA hybrid, a DNA-protein complex or an RNA-protein complex.

A further aspect of the invention relates to an immunological composition that when introduced into an individual, preferably a human, capable of having induced within it an immunological response, induces an immunological response in such individual to a BASB132 polynucleotide and/or polypeptide encoded therefrom, wherein the composition comprises a recombinant BASB132 polynucleotide and/or polypeptide encoded therefrom

and/or comprises DNA and/or RNA which encodes and expresses an antigen of said BASB132 polynucleotide, polypeptide encoded therefrom, or other polypeptide of the invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity and/or cellular immunity, such as cellular immunity arising from CTL or CD4+ T cells.

A BASB132 polypeptide or a fragment thereof may be fused with co-protein or chemical moiety which may or may not by itself produce antibodies, but which is capable of stabilizing the first protein and producing a fused or modified protein which will have antigenic and/or immunogenic properties, and preferably protective properties. Thus fused recombinant protein, preferably further comprises an antigenic co-protein, such as lipoprotein D from *Haemophilus influenzae*, Glutathione-S-transferase (GST) or beta-galactosidase, or any other relatively large co-protein which solubilizes the protein and facilitates production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system of the organism receiving the protein. The co-protein may be attached to either the amino- or carboxy-terminus of the first protein.

In a vaccine composition according to the invention, a BASB132 polypeptide and/or polynucleotide, or a fragment, or a mimotope, or a variant thereof may be present in a vector, such as the live recombinant vectors described above for example live bacterial vectors.

Also suitable are non-live vectors for the BASB132 polypeptide, for example bacterial outer-membrane vesicles or "blebs". OM blebs are derived from the outer membrane of the two-layer membrane of Gram-negative bacteria and have been documented in many Gram-negative bacteria (Zhou, L *et al.* 1998. *FEMS Microbiol. Lett.* 163:223-228) including *C. trachomatis* and *C. psittaci*. A non-exhaustive list of bacterial pathogens reported to produce blebs also includes: *Bordetella pertussis*, *Borrelia burgdorferi*,

Brucella melitensis, Brucella ovis, Escherichia coli, Haemophilus influenza, Legionella pneumophila, Moraxella catarrhalis, Neisseria gonorrhoeae, Neisseria meningitidis, Pseudomonas aeruginosa and Yersinia enterocolitica.

Blebs have the advantage of providing outer-membrane proteins in their native conformation and are thus particularly useful for vaccines. Blebs can also be improved for vaccine use by engineering the bacterium so as to modify the expression of one or more molecules at the outer membrane. Thus for example the expression of a desired immunogenic protein at the outer membrane, such as the BASB132 polypeptide, can be introduced or upregulated (e.g. by altering the promoter). Instead or in addition, the expression of outer-membrane molecules which are either not relevant (e.g. unprotective antigens or immunodominant but variable proteins) or detrimental (e.g. toxic molecules such as LPS, or potential inducers of an autoimmune response) can be downregulated. These approaches are discussed in more detail below.

The non-coding flanking regions of the BASB132 gene contain regulatory elements important in the expression of the gene. This regulation takes place both at the transcriptional and translational level. The sequence of these regions, either upstream or downstream of the open reading frame of the gene, can be obtained by DNA sequencing. This sequence information allows the determination of potential regulatory motifs such as the different promoter elements, terminator sequences, inducible sequence elements, repressors, elements responsible for phase variation, the shine-dalgarno sequence, regions with potential secondary structure involved in regulation, as well as other types of regulatory motifs or sequences. This sequence is a further aspect of the invention.

This sequence information allows the modulation of the natural expression of the BASB132 gene. The upregulation of the gene expression may be accomplished by altering the promoter, the shine-dalgarno sequence, potential repressor or operator elements, or any other elements involved. Likewise, downregulation of expression can be

achieved by similar types of modification. Alternatively, by changing phase variation sequences, the expression of the gene can be put under phase variation control, or it may be uncoupled from this regulation. In another approach, the expression of the gene can be put under the control of one or more inducible elements allowing regulated expression. Examples of such regulation include, but are not limited to, induction by temperature shift, addition of inductor substrates like selected carbohydrates or their derivatives, trace elements, vitamins, co-factors, metal ions, etc.

Such modifications as described above can be introduced by several different means. The modification of sequences involved in gene expression can be carried out *in vivo* by random mutagenesis followed by selection for the desired phenotype. Another approach consists in isolating the region of interest and modifying it by random mutagenesis, or site-directed replacement, insertion or deletion mutagenesis. The modified region can then be reintroduced into the bacterial genome by homologous recombination, and the effect on gene expression can be assessed. In another approach, the sequence knowledge of the region of interest can be used to replace or delete all or part of the natural regulatory sequences. In this case, the regulatory region targeted is isolated and modified so as to contain the regulatory elements from another gene, a combination of regulatory elements from different genes, a synthetic regulatory region, or any other regulatory region, or to delete selected parts of the wild-type regulatory sequences. These modified sequences can then be reintroduced into the bacterium via homologous recombination into the genome. A non-exhaustive list of preferred promoters that could be used for up-regulation of gene expression includes the promoters *porA*, *porB*, *lbpB*, *tbpB*, *p110*, *lst*, *hpuAB* from *N. meningitidis* or *N. gonorrhoeae*; *ompCD*, *copB*, *lbpB*, *ompE*, *UspA1*; *UspA2*; *TbpB* from *M. Catarrhalis*; *p1*, *p2*, *p4*, *p5*, *p6*, *lpD*, *tbpB*, *D15*, *Hia*, *Hmw1*, *Hmw2* from *H. influenzae*.

In one example, the expression of the gene can be modulated by exchanging its promoter with a stronger promoter (through isolating the upstream sequence of the gene, *in vitro*

modification of this sequence, and reintroduction into the genome by homologous recombination). Upregulated expression can be obtained in both the bacterium as well as in the outer membrane vesicles shed (or made) from the bacterium.

In other examples, the described approaches can be used to generate recombinant bacterial strains with improved characteristics for vaccine applications. These can be, but are not limited to, attenuated strains, strains with increased expression of selected antigens, strains with knock-outs (or decreased expression) of genes interfering with the immune response, strains with modulated expression of immunodominant proteins, strains with modulated shedding of outer-membrane vesicles.

Thus, also provided by the invention is a modified upstream region of the BASB132 gene, which modified upstream region contains a heterologous regulatory element which alters the expression level of the BASB132 protein located at the outer membrane. The upstream region according to this aspect of the invention includes the sequence upstream of the BASB132 gene. The upstream region starts immediately upstream of the BASB132 gene and continues usually to a position no more than about 1000 bp upstream of the gene from the ATG start codon. In the case of a gene located in a polycistronic sequence (operon) the upstream region can start immediately preceding the gene of interest, or preceding the first gene in the operon. Preferably, a modified upstream region according to this aspect of the invention contains a heterologous promotor at a position between 500 and 700 bp upstream of the ATG.

Thus, the invention provides a BASB132 polypeptide, in a modified bacterial bleb. The invention further provides modified host cells capable of producing the non-live membrane-based bleb vectors. The invention further provides nucleic acid vectors comprising the BASB132 gene having a modified upstream region containing a heterologous regulatory element.

Further provided by the invention are processes to prepare the host cells and bacterial blebs according to the invention.

Also provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides and/or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof, which have been shown to encode non-variable regions of bacterial cell surface proteins, in polynucleotide constructs used in such genetic immunization experiments in animal models of infection with *Moraxella catarrhalis*. Such experiments will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value, derived from the requisite organ of the animal successfully resisting or clearing infection, for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *Moraxella catarrhalis* infection, in mammals, particularly humans.

The invention also includes a vaccine formulation which comprises an immunogenic recombinant polypeptide and/or polynucleotide of the invention together with a suitable carrier, such as a pharmaceutically acceptable carrier. Since the polypeptides and polynucleotides may be broken down in the stomach, each is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostatic compounds and solutes which render the formulation isotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The

formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use.

The vaccine formulation of the invention may also include adjuvant systems for enhancing the immunogenicity of the formulation. Preferably the adjuvant system raises preferentially a TH1 type of response.

An immune response may be broadly distinguished into two extreme categories, being a humoral or cell mediated immune responses (traditionally characterised by antibody and cellular effector mechanisms of protection respectively). These categories of response have been termed TH1-type responses (cell-mediated response), and TH2-type immune responses (humoral response).

Extreme TH1-type immune responses may be characterised by the generation of antigen specific, haplotype restricted cytotoxic T lymphocytes, and natural killer cell responses. In mice TH1-type responses are often characterised by the generation of antibodies of the IgG2a subtype, whilst in the human these correspond to IgG1 type antibodies. TH2-type immune responses are characterised by the generation of a broad range of immunoglobulin isotypes including in mice IgG1, IgA, and IgM.

It can be considered that the driving force behind the development of these two types of immune responses are cytokines. High levels of TH1-type cytokines tend to favour the induction of cell mediated immune responses to the given antigen, whilst high levels of TH2-type cytokines tend to favour the induction of humoral immune responses to the antigen.

The distinction of TH1 and TH2-type immune responses is not absolute. In reality an individual will support an immune response which is described as being predominantly

TH1 or predominantly TH2. However, it is often convenient to consider the families of cytokines in terms of that described in murine CD4 +ve T cell clones by Mosmann and Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, p145-173*). Traditionally, TH1-type responses are associated with the production of the INF- γ and IL-2 cytokines by T-lymphocytes. Other cytokines often directly associated with the induction of TH1-type immune responses are not produced by T-cells, such as IL-12. In contrast, TH2- type responses are associated with the secretion of IL-4, IL-5, IL-6 and IL-13.

It is known that certain vaccine adjuvants are particularly suited to the stimulation of either TH1 or TH2 - type cytokine responses. Traditionally the best indicators of the TH1:TH2 balance of the immune response after a vaccination or infection includes direct measurement of the production of TH1 or TH2 cytokines by T lymphocytes *in vitro* after restimulation with antigen, and/or the measurement of the IgG1:IgG2a ratio of antigen specific antibody responses.

Thus, a TH1-type adjuvant is one which preferentially stimulates isolated T-cell populations to produce high levels of TH1-type cytokines when re-stimulated with antigen *in vitro*, and promotes development of both CD8+ cytotoxic T lymphocytes and antigen specific immunoglobulin responses associated with TH1-type isotype.

Adjuvants which are capable of preferential stimulation of the TH1 cell response are described in International Patent Application No. WO 94/00153 and WO 95/17209.

3 De-O-acylated monophosphoryl lipid A (3D-MPL) is one such adjuvant. This is known from GB 2220211 (Ribi). Chemically it is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains and is manufactured by Ribi

Immunochem, Montana. A preferred form of 3 De-O-acylated monophosphoryl lipid A is disclosed in European Patent 0 689 454 B1 (SmithKline Beecham Biologicals SA).

Preferably, the particles of 3D-MPL are small enough to be sterile filtered through a 0.22micron membrane (European Patent number 0 689 454).

3D-MPL will be present in the range of 10µg - 100µg preferably 25-50µg per dose wherein the antigen will typically be present in a range 2-50µg per dose.

Another preferred adjuvant comprises QS21, an Hplc purified non-toxic fraction derived from the bark of Quillaja Saponaria Molina. Optionally this may be admixed with 3 De-O-acylated monophosphoryl lipid A (3D-MPL), optionally together with an carrier.

The method of production of QS21 is disclosed in US patent No. 5,057,540.

Non-reactogenic adjuvant formulations containing QS21 have been described previously (WO 96/33739). Such formulations comprising QS21 and cholesterol have been shown to be successful TH1 stimulating adjuvants when formulated together with an antigen.

Further adjuvants which are preferential stimulators of TH1 cell response include immunomodulatory oligonucleotides, for example unmethylated CpG sequences as disclosed in WO 96/02555.

Combinations of different TH1 stimulating adjuvants, such as those mentioned hereinabove, are also contemplated as providing an adjuvant which is a preferential stimulator of TH1 cell response. For example, QS21 can be formulated together with 3D-MPL. The ratio of QS21 : 3D-MPL will typically be in the order of 1 : 10 to 10 : 1; preferably 1:5 to 5 : 1 and often substantially 1 : 1. The preferred range for optimal synergy is 2.5 : 1 to 1 : 1 3D-MPL: QS21.

Preferably a carrier is also present in the vaccine composition according to the invention. The carrier may be an oil in water emulsion, or an aluminium salt, such as aluminium phosphate or aluminium hydroxide.

A preferred oil-in-water emulsion comprises a metabolisable oil, such as squalene, alpha tocopherol and Tween 80. In a particularly preferred aspect the antigens in the vaccine composition according to the invention are combined with QS21 and 3D-MPL in such an emulsion. Additionally the oil in water emulsion may contain span 85 and/or lecithin and/or tricaprylin.

Typically for human administration QS21 and 3D-MPL will be present in a vaccine in the range of 1µg - 200µg, such as 10-100µg, preferably 10µg - 50µg per dose.

Typically the oil in water will comprise from 2 to 10% squalene, from 2 to 10% alpha tocopherol and from 0.3 to 3% tween 80. Preferably the ratio of squalene: alpha tocopherol is equal to or less than 1 as this provides a more stable emulsion. Span 85 may also be present at a level of 1%. In some cases it may be advantageous that the vaccines of the present invention will further contain a stabiliser.

Non-toxic oil in water emulsions preferably contain a non-toxic oil, e.g. squalane or squalene, an emulsifier, e.g. Tween 80, in an aqueous carrier. The aqueous carrier may be, for example, phosphate buffered saline.

A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil in water emulsion is described in WO 95/17210.

The present invention also provides a polyvalent vaccine composition comprising a vaccine formulation of the invention in combination with other antigens, in particular antigens useful for treating cancers, autoimmune diseases and related conditions. Such a

polyvalent vaccine composition may include a TH-1 inducing adjuvant as hereinbefore described.

While the invention has been described with reference to certain BASB132 polypeptides and polynucleotides, it is to be understood that this covers fragments of the naturally occurring polypeptides and polynucleotides, and similar polypeptides and polynucleotides with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant polypeptides or polynucleotides.

Compositions, kits and administration

In a further aspect of the invention there are provided compositions comprising a BASB132 polynucleotide and/or a BASB132 polypeptide for administration to a cell or to a multicellular organism.

The invention also relates to compositions comprising a polynucleotide and/or a polypeptides discussed herein or their agonists or antagonists. The polypeptides and polynucleotides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to an individual. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide and/or polynucleotide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides, polynucleotides and other compounds of the invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide and/or polynucleotide, such as the soluble form of a polypeptide and/or polynucleotide of the present invention, agonist or antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Polypeptides, polynucleotides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other compounds of the present invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, solutions, powders and the like.

For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 $\mu\text{g/kg}$ of subject.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.5-5 microgram/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Sequence Databases, Sequences in a Tangible Medium, and Algorithms

Polynucleotide and polypeptide sequences form a valuable information resource with which to determine their 2- and 3-dimensional structures as well as to identify further sequences of similar homology. These approaches are most easily facilitated by storing the sequence in a computer readable medium and then using the stored data in a known macromolecular structure program or to search a sequence database using well known searching tools, such as the GCG program package.

Also provided by the invention are methods for the analysis of character sequences or strings, particularly genetic sequences or encoded protein sequences. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such as identity and similarity analysis, DNA, RNA and protein structure analysis, sequence assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, codon usage analysis, nucleic acid base trimming, and sequencing chromatogram peak analysis.

A computer based method is provided for performing homology identification. This method comprises the steps of: providing a first polynucleotide sequence comprising the sequence of a polynucleotide of the invention in a computer readable medium; and comparing said first polynucleotide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

A computer based method is also provided for performing homology identification, said method comprising the steps of: providing a first polypeptide sequence comprising the sequence of a polypeptide of the invention in a computer readable medium; and comparing said first polypeptide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their

entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

DEFINITIONS

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heine, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GAP program in the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN (Altschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990), and FASTA(Pearson and Lipman Proc. Natl. Acad. Sci. USA 85; 2444-2448 (1988). The BLAST family of programs is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894;

Altschul, S., *et al.*, *J. Mol. Biol.* 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

Parameters for polypeptide sequence comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: BLOSSUM62 from Henikoff and Henikoff,
Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992)

Gap Penalty: 8

Gap Length Penalty: 2

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

Parameters for polynucleotide comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

Available as: The "gap" program from Genetics Computer Group, Madison WI. These are the default parameters for nucleic acid comparisons.

A preferred meaning for "identity" for polynucleotides and polypeptides, as the case may be, are provided in (1) and (2) below.

- (1) Polynucleotide embodiments further include an isolated polynucleotide comprising a polynucleotide sequence having at least a 50, 60, 70, 80, 85, 90, 95, 97 or 100% identity to the reference sequence of SEQ ID NO:1, wherein said polynucleotide sequence may be identical to the reference sequence of SEQ ID NO:1 or may include up to a certain integer number of nucleotide alterations as compared to the reference

sequence, wherein said alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of nucleotide alterations, x_n is the total number of nucleotides in SEQ ID NO:1, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1, that is it may be 100% identical, or it may include up to a certain integer number of nucleic acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one nucleic acid deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between those terminal positions, interspersed either individually

among the nucleic acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleic acid alterations for a given percent identity is determined by multiplying the total number of nucleic acids in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleic acids in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of nucleic acid alterations, x_n is the total number of nucleic acids in SEQ ID NO:1, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n .

(2) Polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 50,60, 70, 80, 85, 90, 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:2, wherein said polypeptide sequence may be identical to the reference sequence of SEQ ID NO:2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of amino acid alterations is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

By way of example, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is it may be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

"Individual(s)," when used herein with reference to an organism, means a multicellular eukaryote, including, but not limited to a metazoan, a mammal, an ovid, a bovid, a simian, a primate, and a human.

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA including single and double-stranded regions.

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino

acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Disease(s)" means any disease caused by or related to infection by a bacteria, including, for example, otitis media in infants and children, pneumonia in elderlies, sinusitis, nosocomial infections and invasive diseases, chronic otitis media with hearing loss, fluid accumulation in the middle ear, auditive nerve damage, delayed speech learning, infection of the upper respiratory tract and inflammation of the middle ear.

EXAMPLES:

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples are illustrative, but do not limit the invention.

Example 1: DNA sequencing of the BASB132 gene from *Moraxella catarrhalis* strain ATCC 43617.

A: BASB132 in *Moraxella catarrhalis* strain.

The DNA sequence of the BASB132 gene from the *Moraxella catarrhalis* strain ATCC 43617 (also referred to as strain MC2931) is shown in SEQ ID NO:1. The translation of the BASB132 polynucleotide sequence showed in SEQ ID NO:2.

B: BASB132 in *Moraxella catarrhalis* strain 43617.

The sequence of the BASB132 gene was confirmed in *Moraxella catarrhalis* strain ATCC 43617. For this purpose, plasmid DNA (see example 2A) containing the gene region encoding the truncated BASB132 from *Moraxella catarrhalis* strain ATCC 43617 (corresponding to the last 2977 nucleotides 3' end) was submitted to DNA sequencing using the Big Dyes kit (Applied biosystems) and analyzed on a ABI 373/A DNA sequencer in the conditions described by the supplier using primers *Moraxella catarrhalis* oli 5 YtfN (5'-CCC ACC GTT TGC CCT TCA AT-3') [SEQ ID NO:5] and oli 6 YtfN (5'-CAT TTT TCC CGA GCA TTC AAA C-3') [SEQ ID NO:6] and oli 7 YtfN (5'-AAT CAG CCA CTT ATC GCC AC-3') [SEQ ID NO:7] specific for the BASB132 gene and M13 Universal Sequence Primer (5'-GTA AAA CGA CGG CCA GT-3') [SEQ ID NO:8] and M13 Reverse Sequence Primer (5'-CAG GAA ACA GCT ATG AC-3') [SEQ ID NO:9] specific for the vector. As a result, the polynucleotide and deduced polypeptide sequences, referred to as SEQ ID NO:3 and SEQ ID NO:4 respectively, were obtained.

Using the MegAlign program from the DNASTAR software package, an alignment of the polynucleotide sequences of SEQ ID NO:1 and 3 was performed, and is displayed in Figure 1; a pairwise comparison of identities shows that the two BASB132 polynucleotide gene sequences are 100% identical in the overlapping region. Using the same MegAlign program, an alignment of the polypeptide sequences of SEQ ID NO:2 and 4 was performed, and is displayed in Figure 2; a pairwise comparison of identities shows that the two BASB132 protein sequences are 100% identical in the overlapping region.

Example 2: Construction of Plasmid to Express Recombinant truncated BASB132

A: Cloning of truncated BASB132.

The *Bsp*HI and *Bgl*II restriction sites engineered into the oli1 YtfN (5'-TCA TGA ATG ACT CAG GCA AAG-3') [SEQ ID NO:10] forward and oli2 YtfN (5'-AGA TCT AAA CTT CCA ACG ATA AAT C-3') [SEQ ID NO:11] reverse amplification primers, respectively, permitted directional cloning of the PCR product into the *E.coli* expression plasmid pQE60 such that a truncated BASB132 protein could be expressed as a fusion protein containing a (His)₆ affinity chromatography tag at the C-terminus. Due to the large size of the gene, two parallel PCR experiments were carried out to amplify the entire coding sequence using oli1 YtfN (5'-TCA TGA ATG ACT CAG GCA AAG-3') [SEQ ID NO:10] and oli3 (5'-AAA CAA ATC GCA CCC ACG CC-3') [SEQ ID NO:12] for one amplification, and oli2 YtfN (5'-AGA TCT AAA CTT CCA ACG ATA AAT C-3') [SEQ ID NO:11] and oli4 YtfN (5'-ACA AAT TGC AGC GCA TTG TTG G-3') [SEQ ID NO:13] for the second amplification. Both amplified fragments share a common region with a *Bst*XI single restriction site. The BASB132 PCR products were first introduced into the pCRII TOPO cloning vector (Invitrogen) using Top10 bacterial cells, according to the manufacturer's instructions. This intermediate construct was realized to facilitate further cloning into an expression vector. Transformants containing the BASB132 DNA insert were selected by restriction

enzyme analysis. Following digestion, a ~20µl aliquot of the reaction was analyzed by agarose gel electrophoresis (0.8 % agarose in a Tris-acetate-EDTA (TAE) buffer). DNA fragments were visualized by UV illumination after gel electrophoresis and ethidium bromide staining. A DNA molecular size standard (1 Kb ladder, Life Technologies) was electrophoresed in parallel with the test samples and was used to estimate the size of the DNA fragments. Plasmid purified from selected transformants for each cloning were then sequentially digested to completion with *Bsp*HI and *Bst*XI or *Bst*XI and *Bg*/II restriction enzymes as recommended by the manufacturer (Life Technologies). The digested DNA fragment was then purified using silica gel-based spin columns prior to ligation with the pQE60 plasmid.

B: Expression Analysis of PCR-Positive Transformants.

To prepare the expression plasmid pQE60 for ligation, it was similarly digested to completion with both *Nco*I and *Bg*/II and then treated with calf intestinal phosphatase (CIP, ~0.02 units / pmol of 5' end, Life Technologies) as directed by the manufacturer to prevent self ligation. An approximately 5-fold molar excess of the digested fragments to the prepared vector was used to program the ligation reaction. A standard ~20 µl ligation reaction (~16°C, ~16 hours), using methods well known in the art, was performed using T4 DNA ligase (~2.0 units / reaction, Life Technologies). An aliquot of the ligation (~5 µl) was used to transform electro-competent cells according to methods well known in the art. Following a ~2-3 hour outgrowth period at 37°C in ~1.0 ml of LB broth, transformed cells were plated on LB agar plates containing ampicillin (100 µg/ml). Antibiotic was included in the selection. Plates were incubated overnight at 37°C for ~16 hours. Individual ApR colonies were picked with sterile toothpicks and used to "patch" inoculate fresh LB ApR plates as well as a ~1.0 ml LB ApR broth culture. Both the patch plates and the broth culture were incubated overnight at 37°C in either a standard incubator (plates) or a shaking water bath. Restriction analysis was then performed to verify that transformants contained the BASB132 DNA insert. Following digestion, a ~20µl aliquot of the reaction was analyzed by agarose gel

electrophoresis (0.8 % agarose in a Tris-acetate-EDTA (TAE) buffer). DNA fragments were visualized by UV illumination after gel electrophoresis and ethidium bromide staining. A DNA molecular size standard (1 Kb ladder, Life Technologies) was electrophoresed in parallel with the test samples and was used to estimate the size of the DNA fragments. Transformants that produced the expected size DNA fragment were identified as strains containing a BASB132 expression construct. Expression plasmid containing strains were then analyzed for the inducible expression of recombinant BASB132.

C: Expression Analysis of PCR-Positive Transformants.

An aliquot (~1 µl) of the recombinant plasmid DNA preparations were then transformed into competent M15(pREP4) bacterial cells according to methods well known in the art. Following a ~2-3 hours outgrowth period at 37°C in ~1.0 ml of LB broth, transformed cells were plated on LB agar plates containing ampicillin (100 µg/ml) and kanamycin (30 µg/ml). Antibiotic was included in the selection. Plates were incubated overnight at 37°C for ~16 hours. Individual Ap^R Km^R colonies were picked with sterile toothpicks and used to inoculate ~5.0 ml LB Ap^R Km^R broth culture. The broth cultures were incubated overnight at 37°C with shaking (~250 rpm). An aliquot of the overnight seed culture (~1.0 ml) was inoculated into a 125 ml erlenmeyer flask containing ~25 ml of LB Ap broth and grown at 37 °C with shaking (~250 rpm) until the culture turbidity reached O.D.600 of ~0.5, i.e. mid-log phase (usually about 1.5 - 2.0 hours). At this time approximately half of the culture (~12.5 ml) was transferred to a second 125 ml flask and expression of recombinant BASB132 protein induced by the addition of IPTG (1.0 M stock prepared in sterile water, Sigma) to a final concentration of 1.0 mM. Incubation of both the IPTG-induced and non-induced cultures continued for an additional ~4 hours at 37 °C with shaking. Samples (~1.0 ml) of both induced and non-induced cultures were removed after the induction period and the cells collected by centrifugation in a microcentrifuge at room temperature for ~3 minutes. Individual cell pellets were suspended in ~50µl of sterile water, then mixed with an equal volume of

2X Laemelli SDS-PAGE sample buffer containing 2-mercaptoethanol, and placed in boiling water bath for ~3 min to denature protein. Equal volumes (~15 µl) of both the crude IPTG-induced and the non-induced cell lysates were loaded onto duplicate 12% Tris/glycine polyacrylamide gel (1 mm thick Mini-gels, Novex). The induced and non-induced lysate samples were electrophoresed together with prestained molecular weight markers (SeeBlue, Novex) under conventional conditions using a standard SDS/Tris/glycine running buffer (BioRad). Following electrophoresis, one gel was stained with commassie brilliant blue R250 (BioRad) and then destained to visualize novel BASB132 IPTG-inducible protein(s) (Figure 3A). The second gel was electroblotted onto a PVDF membrane (0.45 micron pore size, Novex) for ~2 hrs at 4 °C using a BioRad Mini-Protean II blotting apparatus and Towbin's methanol (20 %) transfer buffer. Blocking of the membrane and antibody incubations were performed according to methods well known in the art. A monoclonal anti-RGS (His)₃ antibody, followed by a second rabbit anti-mouse antibody conjugated to HRP (QiaGen), was used to confirm the expression and identity of the BASB132 recombinant protein (Figure 3B). Visualization of the anti-His antibody reactive pattern was achieved using either an ABT insoluble substrate or using Hyperfilm with the Amersham ECL chemiluminescence system.

Example 3: Production of Recombinant BASB132

Bacterial strain

A recombinant expression strain of *E. coli* M15(PREP4) containing a plasmid (pQE60) encoding BASB132 from *M. catarrhalis*, was used to produce cell mass for purification of recombinant protein. The expression strain was cultivated on LB agar plates containing 100 µg/ml ampicillin ("Ap") and 30 µg/ml kanamycin ("Km") to ensure that pQE60 and pREP4 were maintained. For cryopreservation at -80 °C, the strain was propagated in LB broth containing the same concentration of antibiotics then mixed with an equal volume of LB broth containing 30% (w/v) glycerol.

Media

The fermentation medium used for the production of recombinant protein consisted of 2X YT broth (Difco) containing 100 µg/ml Ap and 30 µg/ml Km. Antifoam was added to medium for the fermentor at 0.25 ml/L (Antifoam 204, Sigma). To induce expression of the BASB132 recombinant protein, IPTG (Isopropyl β-D-Thiogalactopyranoside) was added to the fermentor (1 mM, final).

Fermentation

A 500-ml erlenmeyer seed flask, containing 50ml working volume, was inoculated with 0.3 ml of rapidly thawed frozen culture, or several colonies from a selective agar plate culture, and incubated for approximately 12 hours at $37 \pm 1^\circ\text{C}$ on a shaking platform at 150rpm (Innova 2100, New Brunswick Scientific). This seed culture was then used to inoculate a 5-L working volume fermentor containing 2X YT broth and both Ap antibiotics. The fermentor (Bioflo 3000, New Brunswick Scientific) was operated at $37 \pm 1^\circ\text{C}$, 0.2 - 0.4 VVM air sparge, 250 rpm in Rushton impellers. The pH was not controlled in either the flask seed culture or the fermentor. During fermentation, the pH ranged 6.5 to 7.3 in the fermentor. IPTG (1.0 M stock, prepared in sterile water) was added to the fermentor when the culture reached mid-log of growth (~ 0.7 O.D.600 units). Cells were induced for 2 - 4 hours then harvested by centrifugation using either a 28RS Heraeus (Sepatech) or RC5C superspeed centrifuge (Sorvall Instruments). Cell paste was stored at -20°C until processed.

Chemicals and Materials.

Imidazole and Triton X-100 were purchased from Merck. Guanidine hydrochloride was from Fluka. Aprotinin was obtained from Sigma Chemical Company. Urea and AEBSF were from ICN-Biochemicals. All other chemicals were reagent grade or better.

Ni-NTA Superflow resin and Penta-His Antibody, BSA free were obtained from QiaGen. MicroBCA assay was obtained from Pierce; Amicon 3 filters from Millipore. Dialysis membrane (MWCO12-14000) were from MFPI, USA. Molecular mass marker (BenchMark ladder) was from Life-technologies.

Example 4: Purification of recombinant BASB132 from *E. coli*.

Extraction-Purification

Cell paste from 2750 ml IPTG induced culture (~4 hours, OD₆₂₀= 0.5) was resuspended in 110 ml of 100 mM NaH₂PO₄, 10mM Tris-HCl buffer pH 8 containing 6M Guanidium Chloride (Buffer A) and left for 1 hour at room temperature and then was centrifuged at 10,000g for 20 minutes. Supernatant was applied on Ni-NTA superflow resin equilibrated in buffer A and was immediately collected. Resin was washed twice with 100 mM NaH₂PO₄, 10 mM Tris-HCl buffer pH 6.3, containing 8M Urea (buffer C). Elution was performed with 4x1.5ml 100mM NaH₂PO₄, 10 mM Tris-HCl buffer pH 5.9, containing 8M Urea (buffer D) followed by 4x1.5 ml 100 mM NaH₂PO₄, 10mM Tris-HCl buffer pH 4.5 containing 8M Urea (buffer E). Fractions were neutralised with 25% volume of 200 mM phosphate buffer pH 7.5. Fractions containing BASB132 protein were pooled and dialyzed against 100 mM NaH₂PO₄ buffer pH 7.4 containing 8M Urea, then 4M Urea, then 2M Urea and finally three times against PBS buffer pH 7.4 containing 0.1% Triton-X100.

Due to precipitation during the dialysis step, the solution was cloudy (solution A). After two centrifugation steps (10 minutes at 3000 rpm), the solution became clear (solution B).

Solutions A and B of purified BASB132 protein were quantified using Micro BCA assay reagent. Solution A was quantified at 170µg/ml. 1.65 mg of purified and clarified protein (solution B) were obtained, at a final concentration of 75µg/ml.

As shown in figure 4-A, purified BASB132 protein appeared in SDS-PAGE analysis as a major band migrating at around 100 kDa (estimated relative molecular mass). Purity was estimated to more than 70 %. BASB132 protein was reactive against a mouse

monoclonal antibody raised against the 6-Histidine motif (figure 4-B).

Example 5: Production of Antisera to Recombinant BASB132

Polyvalent antisera directed against the BASB132 protein are generated by vaccinating rabbits with the purified recombinant BASB132 protein.

Polyvalent antisera directed against the BASB132 protein are also generated by vaccinating mice with the purified recombinant BASB132 protein.

Animals are bled prior to the first immunization ("pre-bleed") and after the last immunization.

Anti-BASB132 protein titres are measured by an ELISA using purified recombinant BASB132 protein. The titre is defined as mid-point titers calculated by 4-parameter logistic model using the XL Fit software.

The antisera are also used as the first antibody to identify the protein in a western blot as described in example 7 below. The western-blot can show the presence of anti-BASB132 antibody in the sera of immunized animals.

Example 6: Immunological characterization: Surface exposure of BASB132

Anti-BASB132 protein titres are determined by an ELISA using formalin-killed whole cells of *Moraxella catarrhalis*. The titre is defined as mid-point titers calculated by 4-parameter logistic model using the XL Fit software.

Example 7. Immunological Characterisation: Western Blot Analysis

Several strains of *M. catarrhalis* including ATCC 43617, as well as clinical isolates from various geographic regions, are grown on Muller Hinton agar plates for 24 hours at 36°C. Several colonies are used to inoculate broth. Cultures are grown until the A₆₂₀ is approximately 0.6 and cells are collected by centrifugation. Cells are then concentrated and solubilized in PAGE sample buffer. The solubilized cells are then

resolved on 4-20% polyacrylamide gels and the separated proteins are electrophoretically transferred to PVDF membranes. The PVDF membranes are then pretreated with saturation buffer. All subsequent incubations are carried out using this pretreatment buffer.

PVDF membranes are incubated with preimmune serum or rabbit or mouse immune seru. PVDF membranes are then washed.

PVDF membranes are incubated with biotin-labeled sheep anti-rabbit or mouse Ig. PVDF membranes are then washed 3 times with wash buffer, and incubated with streptavidin-peroxydase. PVDF membranes are then washed 3 times with wash buffer and developed with 4-chloro-1-naphtol.

Example 8: Immunological characterization: Bactericidal Activity

Complement-mediated cytotoxic activity of anti-BASB132 antibodies is examined to determine the vaccine potential of BASB132 protein antiserum that is prepared as described above. The activities of the pre-immune serum and the anti-BASB132 antiserum in mediating complement killing of *M. catarrhalis* are examined.

Strains of *M. catarrhalis* are grown on plates. Several colonies are added to liquid medium. Cultures are grown and collected until the A620 is approximately 0.4. After one wash step, the pellet is suspended and diluted.

Preimmune sera and the anti-BASB132 sera is deposited into the first well of a 96-wells plate and serial dilutions are deposited in the other wells of the same line. Live diluted *M. catarrhalis* is subsequently added and the mixture is incubated. Complement is added into each well at a working dilution defined beforehand in a toxicity assay.

Each test includes a complement control (wells without serum containing active or inactivated complement source), a positive control (wells containing serum with a know titer of bactericidal antibodies), a culture control (wells without serum and complement) and a serum control (wells without complement).

Bactericidal activity of rabbit or mice antiserum (50% killing of homologous strain) is measured.

Example 9: Presence of Antibody to BASB132 in Human Convalescent Sera

Western blot analysis of purified recombinant BASB132 is performed as described in Example 7 above, except that a pool of human sera from children infected by *M. catarrhalis* is used as the first antibody preparation.

Example 10: Efficacy of BASB132 vaccine: enhancement of lung clearance of *M. catarrhalis* in mice.

This mouse model is based on the analysis of the lung invasion by *M. catarrhalis* following a standard intranasal challenge to vaccinated mice.

Groups of mice are immunized with BASB132 vaccine. After the booster, the mice are challenged by instillation of bacterial suspension into the nostril under anaesthesia.

Mice are killed between 30 minutes and 24 hours after challenge and the lungs are removed aseptically and homogenized individually. The log₁₀ weighted mean number of CFU/lung is determined by counting the colonies grown on agar plates after plating of dilutions of the homogenate. The arithmetic mean of the log₁₀ weighted mean number of CFU/lung and the standard deviations are calculated for each group.

Results are analysed statistically.

In this experiment groups of mice are immunized either with BASB132 or with a killed whole cells (kwc) preparation of *M. catarrhalis* or sham immunized.

Example 11: Inhibition of *M. catarrhalis* adhesion onto cells by anti-BASB132 antiserum.

This assay measures the capacity of anti BASB132 sera to inhibit the adhesion of *Moraxella* bacteria to epithelial cells. This activity could prevent colonization of the nasopharynx by *Moraxella*.

One volume of bacteria is incubated on ice with one volume of pre-immune or anti-BASB132 immune serum dilution. This mixture is subsequently added in the wells of a 24 well plate containing a confluent cells culture that is washed once with culture medium to remove traces of antibiotic. The plate is centrifuged and incubated. Each well is then gently washed. After the last wash, sodium glycocholate is added to the wells. After incubation, the cell layer is scraped and homogenised. Dilutions of the homogenate are plated on agar plates and incubated. The number of colonies on each plate is counted and the number of bacteria present in each well calculated.

Deposited materials

A deposit containing a *Moraxella catarrhalis* Catlin strain has been deposited with the American Type Culture Collection (herein "ATCC") on June 21, 1997 and assigned deposit number 43617. The deposit was described as *Branhamella catarrhalis* (Frosch and Kolle) and is a freeze-dried, 1.5-2.9 kb insert library constructed from *M. catarrhalis* isolate obtained from a transtracheal aspirate of a coal miner with chronic bronchitis. The deposit is described in Antimicrob. Agents Chemother. 21: 506-508 (1982).

The *Moraxella catarrhalis* strain deposit is referred to herein as "the deposited strain" or as "the DNA of the deposited strain."

The deposited strain contains a full length BASB132 gene.

A deposit of the vector pMC-D15 consisting of *Moraxella catarrhalis* DNA inserted in pQE30 has been deposited with the American Type Culture Collection (ATCC) on February 12 1999 and assigned deposit number 207105.

The sequence of the polynucleotides contained in the deposited strain / clone, as well as the amino acid sequence of any polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

The deposit of the deposited strains have been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The deposited strains will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposited strains are provided merely as convenience to those of skill in the art and are not an admission that a deposit is required for enablement, such as that required under 35 U.S.C. §112.

Applicant's or agent's file FB/BM45417 reference number	International application No.
------------------------------------------------------------	-------------------------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 64 lines 1-21.	
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet <input type="checkbox"/>
Name of depositary institution AMERICAN TYPE CULTURE COLLECTION	
Address of depositary institution (including postal code and country) 10801 UNIVERSITY BLVD, MANASSAS, VIRGINIA 20110-2209, UNITED STATES OF AMERICA	
Date of deposit 21 June 1997 and 12 February 1999	Accession Number 43617 and 207105
C. ADDITIONAL INDICATIONS (leave blank if not applicable)	This information is continued on an additional sheet <input type="checkbox"/>
In respect of those designations where a European Patent is sought, a sample of the deposited microorganisms will be made available until the publication of the mention of the grant of the European Patent or until the date on which the application has been refused or withdrawn, only by issue of such a sample to an expert nominated by the person requesting the sample	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input type="checkbox"/>	This sheet was received with the international application
Authorized officer	

For International Bureau use only	
<input type="checkbox"/>	This sheet was received by the International Bureau on:
Authorized officer	

SEQUENCE INFORMATION

BASB132 Polynucleotide and Polypeptide Sequences

SEQ ID NO:1

Moraxella catarrhalis BASB132 polynucleotide sequence from strain ATCC43617

ATGACTCATCAGGATAATTCAAAAACTCGTCATCGGCAGAAAATGGGGTTTCTGCTGGCGTTGCTGCACCGACTAAAA
CTCAGCCCAAACCAACCAACCCACCAACCAACAATGGCTAATCTATTTAATTCGCTTAATTGTCGGAGCAATTATTGC
CGCTTTATTGTTGTTGGCGGTGTTATTTGCCATGACCAACAGCGAGTCAGGTTCAAAGTTTTTAATCGAAAAAATTGCG
TTAGAAACTGGCACTAAGCTTAAGTACAGCGAAGGCTCAATTCGCCATGGAGTTGGGTGCAAGATGTCAAATCGCTC
AAAGTGAAGATATTACCATCACCATTAATCGTGCCTATGTACAGCTTGGGTGGCGAGCCTTGTTTGCTCGCCAAGTGCA
TTTGGTCAATCCTAAGATTGATAAAGTTTATGTGACAAACACCAAGCCATCAACAGGCGAACCCCTTGATTATGCGACC
ATCAACCTACCAAGTGACGCTTAAGCTTGAAAATGCCAAAGTCAATGAAATTATCTATGACCAAGTGGATTCTGAGCCTG
TCGTACTGCATCATATCGCATTTGATCAGCATCATGGGCAGATTCAACAGTTAAAATTGATAACGCCATGCTAAGCTA
TGGTGATGATATTAATATCAGCCATGCCACTGGTGGGAATTGATTAAACAGGTCATTATCCGCTGTCGTTGTCGGCAGAT
GTGCATATTTTGGCACTTGATGATGCGTATTTTGATACTTTGTCGGTGAAAGCAGGCGGTAGCCTTAAGCGTACCGTTG
GTACACTCACTGGCAATATAATCAGCATCATGTGACAGGCGAGCTTTATCGCTCAAGGGTTGGATAAAAACTCACCTTT
TAGCGCACGCCTTGATTTTGATGAAGTACGATTGCCTTATGCTGACAGTCAAAATATTTTACTAAAAAATGGCTCTATC
ATCGCTGATGGCGTCATCTCAAATATCGAGCTACGCATTAACTGAGTTATCCGCCAAAGATATTCTGATGGGCATT
ATCACGGTCGTGGAATTATTCGTGGCAGTACCATGCAATCCCATATTTGACAGGCTGATACTGCCAATGGTACTTTGGT
GGCAACGGGTGATATGACTTGGGAAGATGGCTATGAGCTTGATGCCACCATTACAGCAGACGGCTATCGTATCCGTGAA
GAGATGCCAAGTGATTATCATGAATATAGAGCCTATCTACCTAAGGTTTTGACAGGCTCACTTGGGGTTAAGTATACGC
TATTAGACAAAGCCAGTCATGATACTCGGTTTGAGTTTGATCTGAATCAAAAAGACGGTGAACGCATTCAAGCGACGCT
GGCTCAAAATCAACAGAGTGATCATGAGCCTTGGCGTATTGATGCGACTTGGGCAAATCTAATCCGCCATGATATTCCA
CAAATTGGCGAGATTCTAGCCGCTCAGGTCAGGCATCAGTTCGTTTGGAGAACGGACATACCTATATTAACGCCTCTG
CTGACATTGTTAACTTAATGCCGTCCCAAGCGGATCATATCATGTGCAAGCAAATATTGAGCAAAATCAGCACTTGCA
CTTGACAGATTTTAACATCAAGGTGTGATGGGCGAGTTGACAGGTACTGGCAGGGTTGATTTTGCGACTGCCCAAAGA
CCGCTTGAGTGGCAGCTTGATGTCACTGCCAATCCAGTCAAAACCAATGCGTATTTTCAAACACCCAATCAAACACCGT
TTGAGCAAATTCAGGCAGAATAATTGCTTCAGGGCGTCTGCGTGAGATAGATGGTGTGTCAATTCATGACATTAAAGT
GGATGACAGTGATTTAACCGCTTTATTAATGACGGCAAGCAAGTACATCTGATGGGTCAAGGCGTCAGCAAGATTAGG
CTACAAGATGGACAAATTAGCCATCTTAAGGCAAATTTTGATGGCAAGCTTGCTCAAGATATCTTACCGCAGGTAGCTG
ACTCAAGCATTGGATTGGATATTGAAGGTGATTTGAATGATTTGACCATCACACGGGCAGTCATGATGAATGACTCAGG
CAAAGTATCGCTTGCAAGTCTGCTTGATTAAATGACGGTATTCATTGGCAAATTCAGGGCGATTAGATGAGGTGGAT
ACCGCCGCTTTTGTTGATGATGACAATTTAGTTGCGATCATCACAGGCGATCTTGCCACATCAGGCAGATATCGTGATG
GCGTGATGGATGATATCGCACTCACTTTTGATGGTCAAGTGATTAACAATTCATCCCAAATGGTCATCTAAGTATTGA
TGCGACAGGTTCTGGCAATCGCTTTATGGTCAATCATCTAAGACATGATGGCACAGCAGGCACGCTGAATGCGACAGGC
TGGGTAGATATTAGCCAAGGTGCTGCATGGCAGCTAGAGGCTGATATGAATTCATTGAATTTGGGTGCATTTATCCAGT
CGCTAGATACTGACTTAAACGGTACAATTCAGCTTGCAAGCAACTGGCAAGAGTCTCATCAAGTCATTGATATCAGTAA
TTTGATATTACAGGTAGCTACAATAATCAGCCACTTATCGCCACAGGCAGCTTATATGCCAAGCTTTCCCTACCTAAA
GATTTGGCTGGGTATATCCAAAGTATCAAGCAAGCAAGCCGCCACCGACTTCGTCTGATGATTTATTGGCACTAAAAA
ATCGTATTGATAACAACGCACGCCAAACCCAAAATATCGTTCATAAACTCAATGCTGATAATCTACAAGTCCGCATTGG
CAATAATCATTGGCCATATCAGGCGATGAAAGACAGTTGACCGCCAGTGTGAATGTGATTGATCTTGGGCAGCTGATT
GATACCGCAAGTGGTGGCATTCAAGGCGGTGTGATTGTCTGTCATGATCATCAGCCTTGCCAACTTTATATATTGATG
CCAGTGTCTCATCACTCAGATTTGCCAATATCACCATCCAAAACGCCAAGCTATCGGTAAAATTGTCAATCTGGGTAA
TAGCGAAAGCCAGCTGTTGGTTCAAGGTGATGATATTATCGTGATGGGACGCACCATTAATCTGCTCGGATGGATTTT
AGTGGCACAGAAGCCGATCATATCTTATCAATTTCAACCAAAAAGCGGTGATATCGAAGCATCTATGCATATTGATGGGG
CATTTAATCGCAACAATATGCGATATCATGGTGTGTTGTCGTGATGGTTTTGTCAAGAGTGGATTTGGCAAAATGTCACA
GCGTCAGCCGACCGAGTTAGCTATGGATTGGATGATAATAGCTTACAAATTGCAGCGCATTGTTGGCAATCGGCACAT

ATCCAAGGCGATGGCGTGGGTGCGATTTGTTTACAAGATACATTAAGTTATACACCGCAATCAGGGAATGTCAATCTTA
 TTATCCAAAATCTTGATACTCAAGTACTATTGGCGGCTCTACCCAGTGATATTCATTGGAAATCCATGCTCAATGGTCG
 CATTAAAGGCAATGTGGCAGGCAGGTCAATCACCTTTGGTTGATGCTGTACTGTATTCTGATGATGGTACAATTGGCTTA
 ACTCAAGAAGAGACAGGTTATATTGAGATGCCATATCAGCGTGCGTCTGTGATTGCCAAAAGTGTGTGATAATGGCTTAA
 AGGTAAGAACCAGATATCTTGGGTACGGCAGGTCTGTGGTTATGCCGATGTAATCATCAACCCCAAGCAAACAGACAAGCC
 CATTTCTGGTGCGCTTGTATGAATGATCTTAATTTGGCGGTATTGCGACCATTTTTCCCGAGCATTCAAACATTGAGC
 GGTAAAGTCAGCTTGGCAGGTGGCTTAGGTGGTACATTATCCAAACCATTGTTTTATGGTAATGCCAATTTAAAGAATG
 GTGCATTAGCGATTGTTGGCGTACCAATGCCCATTTCTGATGTTGATGCCACGGTGAACATACGAGGTACTAAAGCACG
 CTTAGACGGTAGATTTATGGGTGGTGGGTGAGGTCATGGTGTGCTTTATGGTGAATGGATTGGGCAGAAGAGCTGTATGCT
 CGCCTTGGCGTATTTGGTGAAAATCTGACCGTCAGCCAGCCACCGCTTGTGACGGCACAAATCAGCCCTGAGCTGGAAG
 TGATTATCAGACCTTTCCAAAATTTGTAGATGTTCAAGGCGTGTGTAGTATCCCATCAGCAACCATTGCTCCGCCAGA
 GGCAACCGCAGATATCGTAACCGAATCCCGAGATGTATCAGTCATTGACCGCGTATCACAGGTAATATTGACCAGATT
 TTAAGTAGAGCGAAACCTTGGGATATTAAATGCAAAATATCGGTGTTGATTGGGTAAATGAGATTGAATTTCTGGGATTG
 GGGCGGTATTACCATTTGGCAGGCGCGATACATTTGACTCAGTCAGGACAAGGTGCAATGCAAGCTCTTGGCGTGGTTCA
 GGTTCCTAAACGCACAAAAATCGATGTATCGGTCAAATTTGGATTAAATATGCCCAAATTCGTTTTGATGGCGAC
 ATGCTTAATCCACGCTTATCTATTGAGGGTGAAAAACAAATGAAGGGCAAACGGTGGGCGTTCGCATTCTGGAACGG
 CATCAATCCAAACATTACCGTATTTAATGATGCAGGTTTGGATGAATACCAAGCAATGAATGCACCTGGTGACAGGACG
 CATTAGCGAATCAAGTGATTGGGTATCACAGAGCAGGTTTTCGCTCACAAGTAACCAATCACCTTGCTGCTGCTGGC
 TTAAGCTTAGGTTTGTGAGGAACCAGAGATTTAACCAACCAAAATGGTCACGCATTGGTTTTGAGAGTTTGACCATTG
 ATGCCTCGGGCAGCTCAGATGATACCAATGTTAATGTTACAGGCTATATTACACCAGATTTGTATTGCGTTATGGCGT
 GGGTGTGTTTAAATGCTGAATCAACCTTATCCATGCGTTATCAATTGACACGCCGTGTTTATATTGAAGCAACCAGTGCC
 GCTGAAAATATGGTTGATGTGATTTATCGTTGGAAGTTTAG

SEQ ID NO:2

Moraxella catarrhalis BASB132 polypeptide sequence deduced from the polynucleotide of
 SeQ ID NO:1

MTHQDNSKNSSSAENGVSAGVAAPTKTQPKPPNPPKQWLIYLIRLIVGAI IAALLLLAVLFAMTNSSESGSKFLIEKIAL
 ETGTLKRYSEGSIRHGVWVQDVKIAQSEDTITINRAYVQLGWRALFARQVHLVNPBKIDKVYVTNTPSTGEFPHYATIN
 LPVTLKLENAKVNEI IYDQVDSEPVVLLHHIAFDHASWADSTVKIDNAML SYGDDINISHATGGIDLTGHYPLSLADVHI
 LALDDAYFDTLVSKAGGSLKRTVGTLTGKYNQHHVTSFIAQGLDKNSPFSARLDFDEVRLPYADSQNILLKNGSI IADG
 VISNIELRINTELSAKDIPDGHYHGRGIIRGSTMQIPYLQADTANGTLVATGDMTWEDGYELDATITADGYRIREEMPSD
 YHEYRAYLPKVLTSGLGVKYTL LDKASHDTRFEFDLNQKDGERIQATLAQNQQSDHEPWRIDATWANLIRHDI PQIGEIH
 SRSGQASVRLENGHTYINASADIVKLNVP SGSYHVQANIEQNQHLHLTDFNYQGVMGELTGTGRVDFATAQRPLEWQLD
 VTANPVKPNAYFQTPNQTFPEQISGRIIASGRLEIDGVS IHDIKVDDSDLTALLNDGKQVHLMGQGVSKIRLQDQGQISH
 LKANFDGKLAQDILPQVADSSIGLDIEGDLNDLTITRAVMMNDSGKVS LAGRLGLNDGIHWQIQGRLEVDTA AFVDDDN
 LVAIITGDLATSGRYRDGVMDIALTFDQVINNSIPNGHLSIDATGSGNRFMVNHLRHDGTAGTLNATGWVDISQGAAW
 QLEADMNSLNLGAFIQSLDLDLNGTIQLAGNWQESHQVIDISNLDITGSYNNQPLIATGSLYAKLSLPKDLAGYIQSIKQ
 ASRPPTSSDDLALKNRIDNNARQTQNI VHKL NADNLQVRIGNNH LAISGDERQLTASVNVIDLGLIDTASGAIQGGVI
 VVNDHHALPTLYIDASVSSLRFANITIQNAQAIGKIVNLGNSESQLLVQGGDIIVMGRTIKSARMDFSGTEADHILSIST
 KSGDIEASMHDGAFNRNRMRYHGVLSDFVKSFGFKMSRQPTFEYSGLDDNSLQIAAHCWQSAHIQGDGVGAICLQDT
 LSYTPQSGNVNLI IQNLDTQVLLAALPSDIHWKSMNLNGRIKAMWQAGQSP LVDVAVLYSDDGTIGLTQEETGYIEMPHYQRA
 SVIAKSVDNGLKVRTDILGTAGRGYADVIINPKQTDKPI SGALVMNDLNLAVLRPFPSIQTLSGKVS LAGGLGGLTSKP
 LFYGNANLKNLALAIIVGVPMPISD VDATVNIRGTRKARLDGRFMGEGHGVLYGEMDWAEEYARLGVFGENLTVSQPPLV
 TAQISPELEVIIRPFQKFDVQGVVSI PSATIRPEATADIVTESPDVSVIDRRITGNIDQILSRAKPWDINANIGVDLG
 NEIEFRGFGAVLPLAGAIHLTQSGQAMQALGVVQVSKRTKIDVIGQNLDLNYAQIRFDGDMNLPRLSIEGEKQIEGQTV
 GVRIRGTASNPNITVFNDAGLDEYQAMNALVTGRISSESSDLGITEQGFRSQVTNHLAAAGLSGLSGTRDLTNQIGHAFG
 FQSLTIDASGSSDDTNVNVTYITPDLYLRGVGVFNAESTLSMRYQLTRRVYIEATSAAENMVDVIYRWKF

SEQ ID NO:3

Moraxella catarrhalis BASB132 polynucleotide sequence from strain ATCC43617

ATGAATGACTCAGGCAAAGTATCGCTTGCAGGTCGCTTGGATTAAATGACGGTATTCATTGGCAAATCAAGGGCGAT
 TAGATGAGGTGGATACCGCCGCTTTTGTGATGATGACAAATTTAGTTGCGATCATCACAGGCGATCTTGCCACATCAGG
 CAGATATCGTGATGGCGTGATGGATGATATCGCACTCACTTTTATGGTCAAGTGATTAACAATTCATCCCAAATGGT
 CATCTAAGTATTGATGCGACAGGTTCTGGCAATCGCTTTTATGGTCAATCATCTAAGACATGATGGCACAGCAGGCACGC
 TGAATGCGACAGGCTGGGTAGATATTAGCCAAGGTGCTGCGATGGCAGCTAGAGGCTGATATGAATTCATTGAATTTGGG
 TGCATTTATCCAGTCGCTAGATACTGACTTAAACGGTACAATTACAGCTTGCAGGCAACTGGCAAGAGTCTCATCAAGTC
 ATTGATATCAGTAATTTGGATATTACAGGTAGCTACAATAATCAGCCACTTATCGCCACAGGCGCTTATATGCCAAGC
 TTTCCCTACCTAAAGATTTGGCTGGGTATATCCAAAGTATCAAGCAAGCAAGCCGCCACCGACTTCGTCTGATGATTT
 ATTTGGCACTAAAAATCGTATTGATAACAACGCACGCCAAACCCAAAATATCGTTCAATAAATCAATGCTGATAATCTA
 CAAGTCCGCATTGGCAATAATCATTGGCCATATCAGGCGATGAAAGACAGTTGACCGCCAGTGTAATGTGATTGATC
 TTGGGCAGCTGATTGATACCGCAAGTGGTGGGATTCAAGGCGGTGTGATTGTGCTCAATGATCATCACGCTTTGCCAAC
 TTTATATATTGATGCCAGTGTCTCATCACTCAGATTTGCCAATATCACCATCCAAACGCCCAAGCTATCGGTAAAATT
 GTCATCTGGGTAATAGCGAAAGCCAGCTGTTGGTTCAAGGTGATGATATTATCGTGATGGGACGCACCATTAATCTG
 CTCGGATGGATTTAGTGGCAGAGCCGATCATATCTTATCAATTTCAACCAAAAGCGGTGATATCGAAGCATCTAT
 GCATATTGATGGGGCATTAAATCGCAACAATATGCGATATCATGGTGTCTTGTCTGATGGTTTTGTCAAGAGTGGATTT
 GGCAAAATGTCACAGCGTCAGCCGACCGAGTTTAGCTATGGATTGGATGATAATAGCTTACAAATTGCAGCGCATTGTT
 GGCAATCGGCACATATCCAAGGCGATGGCGTGGGTGCGATTGTTTACAAGATACATTAAGTTATACACCGCAATCAGG
 GAATGTCAATCTTATTATCCAAAATCTTGATACTCAAGTACTATTGGCGGCTCTACCCAGTGATATTATTGGAATCC
 ATGCTCAATGGTCGATTAAGGCAATGTGGCAGGCGAGTCAATCACCTTTGGTTGATGCTGTACTGTATTCTGATGATG
 GTACAATTGGCTTAATCAAGAAGAGACAGGTTATATTGAGATGCCATATCAGCGTGCGTCTGTGATTGCCAAAAGTGT
 TGATAATGGCTTAAAGTAAGAACCAGATATCTTGGGTACGGCAGGTCGTGGTTATGCCGATGTAATCATCAACCCCAAG
 CAAACAGACAAGCCCATTTCTGGTGGCTTGTATGAATGATCTTAATTTGGCGGTATTGCGACCATTTTTCCCGAGCA
 TTCAAACATTGAGCGGTAAAGTCAGCTTGGCAGGTGGCTTAGGTGGTACATTATCCAAACCATTGTTTTATGGTAATGC
 CAATTTAAAGAAATGGTGCATTAGCGATTGTTGGCGTACCAATGCCCATTTCTGATGTTGATGCCACGGTGAACATACGA
 GGTACTAAAGCAGCCTTAGACGGTAGATTTATGGGTGGTGAGGGTCATGGTGTGCTTTATGGTGAAATGGATTGGGCAG
 AAGAGCTGTATGCTCGCCTTGGCGTATTTGGTGAAAATCTGACCGTCAGCCAGCCACCGCTTGTGACGGCACAAATCAG
 CCCTGAGCTGGAAAGTGATTATCAGACCTTTCCAAAATTTGTAGATGTTCAAGGCGTTGTTAGTATCCCATCAGCAACC
 ATTCGTCCGCGAGAGGCAACCGCAGATATCGTAACCGAATCCCAGATGTATCAGTCATTGACCGCGTATCACAGGTA
 ATATTGACCAGATTTTAAAGTAGAGCGAAACCTTGGGATATTAATGCAAAATATCGGTGTTGATTGGGTAATGAGATTGA
 ATTTCTGTGGATTTGGGGCGGTATTACCATTTGGCAGGCGCGATACATTTGACTCAGTCAGGACAAGGTGCAATGCAAGCT
 CTTGGCGTGGTTAGGTTTCTAAACGCACAAAATCGATGTATCGGTCAAAATTTGGATTAAATTATGCCCAAATTC
 GTTTTGTGATGGGACATGCTTAATCCACGCTTATCTATTGAGGGTGAAAAACAAATTGAAGGGCAACCGGTGGCGTTTCG
 CATTCGTGGAACGGCATCAAATCCAAACATTACCGTATTAAATGATGCAGGTTTGGATGAATACCAAGCAATGAATGCA
 CTGGTGACAGGACGCATTAGCGAATCAAGTATTGGGTATCACAGAGCAGGGTTTTCGCTCACAAGTAACCAATCACC
 TTGCTGCTGCTGGCTTAAGCTTAGGTTTGTGAGGAACAGAGATTTAACCAACCAAAATGGTCACGCATTGGTTTTCA
 GAGTTTGACATTGATGCTCGGGCAGCTCAGATGATACCAATGTTAATGTTACAGGCTATATTACACCAGATTGTAT
 TTGCGTTATGGCGTGGGTGTGTTAATGCTGAATCAACCTTATCCATGCGTTATCAATTGACACGCCGTGTTTATATTG
 AAGCAACCAGTGCCGCTGAAAATATGGTTGATGTGATTATCGTTGGAAGTTTTAG

SEQ ID NO:4

Moraxella catarrhalis BASB132 polypeptide sequence deduced from the polynucleotide of
SeQ ID NO:3

MNDSGKVS LAGRLGLNDGIHWQIQRLDEVDTAAFDVDDNLVAIITGDLATSGRYRDGVMDIALTFDQVINNSIPNGH
 LSIDATGSGNRFMVNHLRHDGTAGTLNATGWVDISQGAAWQLEADMNSLNLGAFIQSLDLDLNGTIQLAGNWQESHQVID

ISNLDITGSYNNQPLIATGSLYARLSLPKDLAGYIQSIKQASRPPTSSD DLLALKNRIDNNARQTQNI VHKL NADNLQVR
IGNNH LAISGDERQLTASVNVIDLGLIDTASGAIQGGVIVVNDHHALPTLYIDASVSSLRFANITIQNAQAIGKIVNLG
NSESQLLVQGGDDIIVMGRTIKSARMDFSGTEADHILSISTKSGDIEASMHDGAFNRNNMRYHGVLS DGFVKSGFGKMSQ
RQPT EFSYGLDDNSLQIAAHCWQSAHIQGDGVGAICLQDTLSYTPQSGNVNLI IQNLDTQVLLAALPSDIHWKSMNLGRI
KAMWQAGQSPLVDAVLYSDDGTIGLTQEETGYIEMPYQRASVIAKSV DNLKVRTDILGTAGRGYADVI INPKQTDKPIS
GALVMNDLNLAVLRPFPSIQTLSGKVS LAGGLGGTLSKPLFYGNANLKN GALAIVGVPMPI SDVDATVNIRG TKARLDG
RFMGGEHGVLYGEMDWAEELYARLGVFGENLTVSQPPLVTAQISPELEVIIRPFQKFVDVQGVVSIP SATIRPPEATAD
IVTESPDVSVIDRRITGNIDQILSRAPWDINANIGVDLNEIEFRGFGAVLPLAGAIHLTQSGQGAMQALGVVQVSKRT
KIDVIGQNLDLNYAQIRFDGDMNPRLSIEGEKQIEGQTVGVRI RGTASNPNITVFNDAGLDEYQAMNALVTGRISSESSD
LGITEQGFRSQVTNHLAAAGLSLGLSGTRDLTNQIGHAFGFQSLTIDASGSSDDTNVNV TGYITPDLYLRYGVGVFNAES
TLSMRYQLTRRVYIEATSA AENMVDVIYRWKF

SEQ ID NO:5**CCC ACC GTT TGC CCT TCA AT****SEQ ID NO:6****CAT TTT TCC CGA GCA TTC AAA C****SEQ ID NO:7****AAT CAG CCA CTT ATC GCC AC****SEQ ID NO:8****GTA AAA CGA CGG CCA GT****SEQ ID NO:9****CAG GAA ACA GCT ATG AC****SEQ ID NO:10****TCA TGA ATG ACT CAG GCA AAG****SEQ ID NO:11****AGA TCT AAA CTT CCA ACG ATA AAT C****SEQ ID NO:12****AAA CAA ATC GCA CCC ACG CC**

SEQ ID NO:13

ACA AAT TGC AGC GCA TTG TTG G

CLAIMS:

1. An isolated polypeptide comprising an amino acid sequence which has at least 85% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4, over the entire length of SEQ ID NO:2 or SEQ ID NO:4 respectively.
2. An isolated polypeptide as claimed in claim 1 in which the amino acid sequence has at least 95% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4, over the entire length of SEQ ID NO:2 or SEQ ID NO:4 respectively.
3. The polypeptide as claimed in claim 1 comprising the amino acid sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4.
4. An isolated polypeptide of SEQ ID NO:2 or SEQ ID NO:4.
5. An immunogenic fragment of the polypeptide as claimed in any one of claims 1 to 4 in which the immunogenic activity of said immunogenic fragment is substantially the same as the polypeptide of SEQ ID NO:2 or SEQ ID NO:4.
6. A polypeptide as claimed in any of claims 1 to 5 wherein said polypeptide is part of a larger fusion protein.
7. An isolated polynucleotide encoding a polypeptide as claimed in any of claims 1 to 6.
8. An isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide that has at least 85% identity to the amino acid sequence of SEQ ID NO:2 or 4 over the

entire length of SEQ ID NO:2 or 4 respectively; or a nucleotide sequence complementary to said isolated polynucleotide.

9. An isolated polynucleotide comprising a nucleotide sequence that has at least 85% identity to a nucleotide sequence encoding a polypeptide of SEQ ID NO:2 or 4 over the entire coding region; or a nucleotide sequence complementary to said isolated polynucleotide.

10. An isolated polynucleotide which comprises a nucleotide sequence which has at least 85% identity to that of SEQ ID NO:1 or 3 over the entire length of SEQ ID NO:1 or 3 respectively; or a nucleotide sequence complementary to said isolated polynucleotide.

11. The isolated polynucleotide as claimed in any one of claims 7 to 10 in which the identity is at least 95% to SEQ ID NO:1 or 3.

12. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:2 or SEQ ID NO:4.

13. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:1 or SEQ ID NO:3.

14. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:2, SEQ ID NO:4 obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or SEQ ID NO:3 or a fragment thereof.

15. An expression vector or a recombinant live microorganism comprising an isolated polynucleotide according to any one of claims 7 - 14.

16. A host cell comprising the expression vector of claim 15 or a subcellular fraction or a membrane of said host cell expressing an isolated polypeptide comprising an amino acid sequence that has at least 85% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4.

17. A process for producing a polypeptide of claims 1 to 6 comprising culturing a host cell of claim 16 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

18. A process for expressing a polynucleotide of any one of claims 7 – 14 comprising transforming a host cell with the expression vector comprising at least one of said polynucleotides and culturing said host cell under conditions sufficient for expression of any one of said polynucleotides.

19. A vaccine composition comprising an effective amount of the polypeptide of any one of claims 1 to 6 and a pharmaceutically acceptable carrier.

20. A vaccine composition comprising an effective amount of the polynucleotide of any one of claims 7 to 14 and a pharmaceutically effective carrier.

21. The vaccine composition according to either one of claims 19 or 20 wherein said composition comprises at least one other *Moraxella catarrhalis* antigen.

22. An antibody immunospecific for the polypeptide or immunological fragment as claimed in any one of claims 1 to 6.

23. A method of diagnosing a *Moraxella catarrhalis* infection, comprising identifying a polypeptide as claimed in any one of claims 1 - 6, or an antibody that is immunospecific

for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

24. Use of a composition comprising an immunologically effective amount of a polypeptide as claimed in any one of claims 1 – 6 in the preparation of a medicament for use in generating an immune response in an animal.

25. Use of a composition comprising an immunologically effective amount of a polynucleotide as claimed in any one of claims 7 - 14 in the preparation of a medicament for use in generating an immune response in an animal.

26. A therapeutic composition useful in treating humans with *Moraxella catarrhalis* disease comprising at least one antibody directed against the polypeptide of claims 1 – 6 and a suitable pharmaceutical carrier.

1/45

Figure 1 : Alignment of the BASB132 polynucleotide sequences.**Identity to SeqID No:1 is indicated by a dot. Gap is indicated by a dash.**

*	20	*
Seqid1	: ATGACTCATCAGGATAATTCAAAAAACTCG	: 30
Seqid3	: -----	: -
	40	60
Seqid1	: TCATCGGCAGAAAATGGGGTTTCTGCTGGC	: 60
Seqid3	: -----	: -
	80	90
Seqid1	: GTTGCTGCACCGACTAAAACTCAGCCCAA	: 90
Seqid3	: -----	: -
	100	120
Seqid1	: CCACCAAACCCACCCAAACCACAATGGCTA	: 120
Seqid3	: -----	: -
	140	150
Seqid1	: ATCTATTTAATTCGCTTAATTGTCGGAGCA	: 150
Seqid3	: -----	: -
	160	180

2/45

Seqid1 : ATTATTGCCGCTTTATTGTTGTTGGCGGTG : 180

Seqid3 : ----- :

* 200 *

Seqid1 : TTATTTGCCATGACCAACAGCGAGTCAGGT : 210

Seqid3 : ----- : -

220 * 240

Seqid1 : TCAAAGTTTTTAATCGAAAAAATTGCGTTA : 240

Seqid3 : ----- : -

* 260 *

Seqid1 : GAAACTGGCACTAAGCTTAAGTACAGCGAA : 270

Seqid3 : ----- : -

280 * 300

Seqid1 : GGCTCAATTCGCCATGGAGTTTGGGTGCAA : 300

Seqid3 : ----- : -

* 320 *

Seqid1 : GATGTCAAAATCGCTCAAAGTGAAGATATT : 330

Seqid3 : ----- : -

3/45

	340	*	360
Seqid1	: ACCATCACCATTAATCGTGCCTATGTACAG	:	360
Seqid3	: -----	:	-
	*	380	*
Seqid1	: CTTGGGTGGCGAGCCTTGTTTGCTCGCCAA	:	390
Seqid3	: -----	:	-
	400	*	420
Seqid1	: GTGCATTTGGTCAATCCTAAGATTGATAAA	:	420
Seqid3	: -----	:	-
	*	440	*
Seqid1	: GTTTATGTGACAAACACCAAGCCATCAACA	:	450
Seqid3	: -----	:	-
	460	*	480
Seqid1	: GGCGAACCCTTTGATTATGCGACCATCAAC	:	480
Seqid3	: -----	:	-
	*	500	*
Seqid1	: CTACCAGTGACGCTTAAGCTTGAAAATGCC	:	510

4/45

Seqid3 : ----- : -

520 * 540

Seqid1 : AAAGTCAATGAAATTATCTATGACCAAGTG : 540

Seqid3 : ----- : -

* 560 *

Seqid1 : GATTCTGAGCCTGTCGTACTGCATCATATC : 570

Seqid3 : ----- : -

580 * 600

Seqid1 : GCATTTGATCACGCATCATGGGCAGATTCA : 600

Seqid3 : ----- : -

* 620 *

Seqid1 : ACAGTTAAAATTGATAACGCCATGCTAAGC : 630

Seqid3 : ----- : -

640 * 660

Seqid1 : TATGGTGATGATATTAATATCAGCCATGCC : 660

Seqid3 : ----- : -

5/45

	*	680	*
Seqid1	:	ACTGGTGGAATTGATTTAACAGGTCATTAT	: 690
Seqid3	:	-----	: -
	700	*	720
Seqid1	:	CCGCTGTCGTTGTCGGCAGATGTGCATATT	: 720
Seqid3	:	-----	: -
	*	740	*
Seqid1	:	TTGGCACTTGATGATGCGTATTTTGATACT	: 750
Seqid3	:	-----	: -
	760	*	780
Seqid1	:	TTGTCGGTGAAAGCAGGCGGTAGCCTTAAG	: 780
Seqid3	:	-----	: -
	*	800	*
Seqid1	:	CGTACCGTTGGTACACTCACTGGCAAATAT	: 810
Seqid3	:	-----	: -
	820	*	840

6/45

Seqid1 : AATCAGCATCATGTGACAGGCAGCTTTATC : 840

Seqid3 : ----- : -

* 860 *

Seqid1 : GCTCAAGGGTTGGATAAAAACTCACCTTTT : 870

Seqid3 : ----- : -

880 * 900

Seqid1 : AGCGCACGCCTTGATTTTGATGAAGTACGA : 900

Seqid3 : ----- : -

* 920 *

Seqid1 : TTGCCTTATGCTGACAGTCAAAATATTTTA : 930

Seqid3 : ----- : -

940 * 960

Seqid1 : CTAAAAAATGGCTCTATCATCGCTGATGGC : 960

Seqid3 : ----- : -

* 980 *

Seqid1 : GTCATCTCAAATATCGAGCTACGCATTAAC : 990

7/45

Seqid3 : ----- : -

1000 * 1020

Seqid1 : ACTGAGTTATCCGCCAAAGATATTCCTGAT : 1020

Seqid3 : ----- : -

* 1040 *

Seqid1 : GGGCATTATCACGGTCGTGGAATTATTCGT : 1050

Seqid3 : ----- : -

1060 * 1080

Seqid1 : GGCAGTACCATGCAAATCCCATATTTGCAG : 1080

Seqid3 : ----- : -

* 1100 *

Seqid1 : GCTGATACTGCCAATGGTACTTTGGTGGCA : 1110

Seqid3 : ----- : -

1120 * 1140

Seqid1 : ACGGGTGATATGACTTGGGAAGATGGCTAT : 1140

Seqid3 : ----- : -

8/45

	*	1160	*
Seqid1	:	GAGCTTGATGCCACCATTACAGCAGACGGC	: 1170
Seqid3	:	-----	: -
		1180	*
Seqid1	:	TATCGTATCCGTGAAGAGATGCCAAGTGAT	: 1200
Seqid3	:	-----	: -
	*	1220	*
Seqid1	:	TATCATGAATATAGAGCCTATCTACCTAAG	: 1230
Seqid3	:	-----	: -
		1240	*
Seqid1	:	GTTTTGACAGGCTCACTTGGGGTTAAGTAT	: 1260
Seqid3	:	-----	: -
	*	1280	*
Seqid1	:	ACGCTATTAGACAAAGCCAGTCATGATACT	: 1290
Seqid3	:	-----	: -
		1300	*
			1320

9/45

Seqid1 : CGGTTTGAGTTTGATCTGAATCAAAAAGAC : 1320

Seqid3 : ----- : -

* 1340 *

Seqid1 : GGTGAACGCATTCAAGCGACGCTGGCTCAA : 1350

Seqid3 : ----- : -

1360 * 1380

Seqid1 : AATCAACAGAGTGATCATGAGCCTTGGCGT : 1380

Seqid3 : ----- : -

* 1400 *

Seqid1 : ATTGATGCGACTTGGGCAAATCTAATCCGC : 1410

Seqid3 : ----- : -

1420 * 1440

Seqid1 : CATGATATTCCACAAATTGGCGAGATTCAT : 1440

Seqid3 : ----- : -

* 1460 *

Seqid1 : AGCCGCTCAGGTCAGGCATCAGTTCGTTTG : 1470

10/45

Seqid3 : ----- : -

1480 * 1500

Seqid1 : GAGAACGGACATACCTATATTAACGCCTCT : 1500

Seqid3 : ----- : -

* 1520 *

Seqid1 : GCTGACATTGTTAAACTTAATGCCGTCCCA : 1530

Seqid3 : ----- : -

1540 * 1560

Seqid1 : AGCGGATCATATCATGTGCAAGCAAATATT : 1560

Seqid3 : ----- : -

* 1580 *

Seqid1 : GAGCAAAATCAGCACTTGCACTTGACAGAT : 1590

Seqid3 : ----- : -

1600 * 1620

Seqid1 : TTTAACTATCAAGGTGTGATGGGCGAGTTG : 1620

Seqid3 : ----- : -

* 1640 *

Seqid1 : ACAGGTACTGGCAGGGTTGATTTTGCGACT : 1650

Seqid3 : ----- : -

1660 * 1680

Seqid1 : GCCCAAAGACCGCTTGAGTGGCAGCTTGAT : 1680

Seqid3 : ----- : -

* 1700 *

Seqid1 : GTCACTGCCAATCCAGTCAAACCCAATGCG : 1710

Seqid3 : ----- : -

1720 * 1740

Seqid1 : TATTTTCAAACACCCAATCAAACACCGTTT : 1740

Seqid3 : ----- : -

* 1760 *

Seqid1 : GAGCAAATTTTCAGGCAGAATAATTGCTTCA : 1770

Seqid3 : ----- : -

1780 * 1800

12/45

Seqid1 : GGGCGTCTGCGTGAGATAGATGGTGTGTCA : 1800

Seqid3 : ----- : -

*

1820

*

Seqid1 : ATTCATGACATTAAAGTGGATGACAGTGAT : 1830

Seqid3 : ----- : -

1840

*

1860

Seqid1 : TTAACCGCTTTATTAAATGACGGCAAGCAA : 1860

Seqid3 : ----- : -

*

1880

*

Seqid1 : GTACATCTGATGGGTCAAGGCGTCAGCAAG : 1890

Seqid3 : ----- : -

1900

*

1920

Seqid1 : ATTAGGCTACAAGATGGACAAATTAGCCAT : 1920

Seqid3 : ----- : -

*

1940

*

Seqid1 : CTTAAGGCAAATTTTGATGGCAAGCTTGCT : 1950

13/45

Seqid3 : ----- : -

1960 * 1980

Seqid1 : CAAGATATCTTACCGCAGGTAGCTGACTCA : 1980

Seqid3 : ----- : -

* 2000 *

Seqid1 : AGCATTGGATTGGATATTGAAGGTGATTTG : 2010

Seqid3 : ----- : -

2020 * 2040

Seqid1 : AATGATTTGACCATCACACGGGCAGTCATG : 2040

Seqid3 : ----- : -

* 2060 *

Seqid1 : ATGAATGACTCAGGCAAAGTATCGCTTGCA : 2070

Seqid3 : : 30

2080 * 2100

Seqid1 : GGTCGTCTTGGATTAAATGACGGTATTCAT : 2100

Seqid3 : : 60

14/45

	*	2120	*
Seqid1	:	TGGCAAATTCAAGGGCGATTAGATGAGGTG	: 2130
Seqid3	:	: 90
		2140	*
Seqid1	:	GATACCGCCGCTTTTGTTGATGATGACAAT	: 2160
Seqid3	:	: 120
	*	2180	*
Seqid1	:	TTAGTTGCGATCATCACAGGCGATCTTGCC	: 2190
Seqid3	:	: 150
		2200	*
Seqid1	:	ACATCAGGCAGATATCGTGATGGCGTGATG	: 2220
Seqid3	:	: 180
	*	2240	*
Seqid1	:	GATGATATCGCACTCACTTTTGATGGTCAA	: 2250
Seqid3	:	: 210
		2260	*
			2280

15/45

Seqid1 : GTGATTAACAATTCAATCCCAAATGGTCAT : 2280

Seqid3 : : 240

*

2300

*

Seqid1 : CTAAGTATTGATGCGACAGGTTCTGGCAAT : 2310

Seqid3 : : 270

2320

*

2340

Seqid1 : CGCTTTATGGTCAATCATCTAAGACATGAT : 2340

Seqid3 : : 300

*

2360

*

Seqid1 : GGCACAGCAGGCACGCTGAATGCGACAGGC : 2370

Seqid3 : : 330

2380

*

2400

Seqid1 : TGGGTAGATATTAGCCAAGGTGCTGCATGG : 2400

Seqid3 : : 360

*

2420

*

Seqid1 : CAGCTAGAGGCTGATATGAATTCATTGAAT : 2430

Seqid3 : : 390

16/45

	2440	*	2460
Seqid1	: TTGGGTGCATTTATCCAGTCGCTAGATACT	:	2460
Seqid3	:	:	420
	*	2480	*
Seqid1	: GACTTAAACGGTACAATTCAGCTTGCAGGC	:	2490
Seqid3	:	:	450
	2500	*	2520
Seqid1	: AACTGGCAAGAGTCTCATCAAGTCATTGAT	:	2520
Seqid3	:	:	480
	*	2540	*
Seqid1	: ATCAGTAATTTGGATATTACAGGTAGCTAC	:	2550
Seqid3	:	:	510
	2560	*	2580
Seqid1	: AATAATCAGCCACTTATCGCCACAGGCAGC	:	2580
Seqid3	:	:	540

17/45

	*	2600	*
Seqid1	:	TTATATGCCAAGCTTTCCTACCTAAAGAT	: 2610
Seqid3	:	: 570
		2620	*
Seqid1	:	TTGGCTGGGTATATCCAAAGTATCAAGCAA	: 2640
Seqid3	:	: 600
	*	2660	*
Seqid1	:	GCAAGCCGCCCACCGACTTCGTCTGATGAT	: 2670
Seqid3	:	: 630
		2680	*
Seqid1	:	TTATTGGCACTAAAAAATCGTATTGATAAC	: 2700
Seqid3	:	: 660
	*	2720	*
Seqid1	:	AACGCACGCCAAACCCAAAATATCGTTCAT	: 2730
Seqid3	:	: 690
		2740	*
Seqid1	:	AAACTCAATGCTGATAATCTACAAGTCCGC	: 2760

18/45

Seqid3 : : 720

* 2780 *

Seqid1 : ATTGGCAATAATCATTTGGCCATATCAGGC : 2790

Seqid3 : : 750

2800 * 2820

Seqid1 : GATGAAAGACAGTTGACCGCCAGTGTGAAT : 2820

Seqid3 : : 780

* 2840 *

Seqid1 : GTGATTGATCTTGGGCAGCTGATTGATAACC : 2850

Seqid3 : : 810

2860 * 2880

Seqid1 : GCAAGTGGTGCGATTCAAGGCGGTGTGATT : 2880

Seqid3 : : 840

* 2900 *

Seqid1 : GTCGTCAATGATCATCACGCTTTGCCAACT : 2910

Seqid3 : : 870

19/45

	2920	*	2940
Seqid1	: TTATATATTGATGCCAGTGTCTCATCACTC	:	2940
Seqid3	:	:	900
	*	2960	*
Seqid1	: AGATTTGCCAATATCACCATCCAAAACGCC	:	2970
Seqid3	:	:	930
	2980	*	3000
Seqid1	: CAAGCTATCGGTAAAATTGTCAATCTGGGT	:	3000
Seqid3	:	:	960
	*	3020	*
Seqid1	: AATAGCGAAAGCCAGCTGTTGGTTCAGGGT	:	3030
Seqid3	:	:	990
	3040	*	3060
Seqid1	: GATGATATTATCGTGATGGGACGCACCATT	:	3060
Seqid3	:	:	1020
	*	3080	*

20/45

Seqid1 : AAATCTGCTCGGATGGATTTTAGTGGCACA : 3090

Seqid3 : : 1050

3100

*

3120

Seqid1 : GAAGCCGATCATATCTTATCAATTTCAACC : 3120

Seqid3 : : 1080

*

3140

*

Seqid1 : AAAAGCGGTGATATCGAAGCATCTATGCAT : 3150

Seqid3 : : 1110

3160

*

3180

Seqid1 : ATTGATGGGGCATTTAATCGCAACAATATG : 3180

Seqid3 : : 1140

*

3200

*

Seqid1 : CGATATCATGGTGTTTTGTCTGATGGTTTT : 3210

Seqid3 : : 1170

3220

*

3240

Seqid1 : GTCAAGAGTGGATTTGGCAAAATGTCACAG : 3240

Seqid3 : : 1200

21/45

	*	3260	*
Seqid1	:	CGTCAGCCGACCGAGTTTAGCTATGGATTG	: 3270
Seqid3	:	: 1230
		3280	*
Seqid1	:	GATGATAATAGCTTACAAATTGCAGCGCAT	: 3300
Seqid3	:	: 1260
	*	3320	*
Seqid1	:	TGTTGGCAATCGGCACATATCCAAGGCGAT	: 3330
Seqid3	:	: 1290
		3340	*
Seqid1	:	GGCGTGGGTGCGATTTGTTTACAAGATACA	: 3360
Seqid3	:	: 1320
	*	3380	*
Seqid1	:	TTAAGTTATACACCGCAATCAGGGAATGTC	: 3390
Seqid3	:	: 1350

22/45

	3400	*	3420
Seqid1	: AATCTTATTATCCAAAATCTTGATACTCAA	:	3420
Seqid3	:	:	1380
	*	3440	*
Seqid1	: GTACTATTGGCGGCTCTACCCAGTGATATT	:	3450
Seqid3	:	:	1410
	3460	*	3480
Seqid1	: CATTGGAAATCCATGCTCAATGGTCGCATT	:	3480
Seqid3	:	:	1440
	*	3500	*
Seqid1	: AAGGCAATGTGGCAGGCAGGTCAATCACCT	:	3510
Seqid3	:	:	1470
	3520	*	3540
Seqid1	: TTGGTTGATGCTGTACTGTATTCTGATGAT	:	3540
Seqid3	:	:	1500
	*	3560	*
Seqid1	: GGTACAATTGGCTTAACTCAAGAAGAGACA	:	3570

23/45

Seqid3 : : 1530

3580 * 3600

Seqid1 : GGTTATATTGAGATGCCATATCAGCGTGCG : 3600

Seqid3 : : 1560

* 3620 *

Seqid1 : TCTGTGATTGCCAAAAGTGTTGATAATGGC : 3630

Seqid3 : : 1590

3640 * 3660

Seqid1 : TTAAAGGTAAGAACCGATATCTTGGGTACG : 3660

Seqid3 : : 1620

* 3680 *

Seqid1 : GCAGGTCGTGGTTATGCCGATGTAATCATC : 3690

Seqid3 : : 1650

3700 * 3720

Seqid1 : AACCCCAAGCAAACAGACAAGCCCATTCT : 3720

Seqid3 : : 1680

24/45

	*	3740	*
Seqid1	:	GGTGCGCTTGTTATGAATGATCTTAATTTG	: 3750
Seqid3	:	: 1710
	3760	*	3780
Seqid1	:	GCGGTATTGCGACCATTTTTCCCGAGCATT	: 3780
Seqid3	:	: 1740
	*	3800	*
Seqid1	:	CAAACATTGAGCGGTAAAGTCAGCTTGGCA	: 3810
Seqid3	:	: 1770
	3820	*	3840
Seqid1	:	GGTGGCTTAGGTGGTACATTATCCAAACCA	: 3840
Seqid3	:	: 1800
	*	3860	*
Seqid1	:	TTGTTTTATGGTAATGCCAATTTAAAGAAT	: 3870
Seqid3	:	: 1830
	3880	*	3900

25/45

Seqid1 : GGTGCATTAGCGATTGTTGGCGTACCAATG : 3900

Seqid3 : : 1860

* 3920 *

Seqid1 : CCCATTTCTGATGTTGATGCCACGGTGAAC : 3930

Seqid3 : : 1890

3940 * 3960

Seqid1 : ATACGAGGTACTAAAGCACGCTTAGACGGT : 3960

Seqid3 : : 1920

* 3980 *

Seqid1 : AGATTTATGGGTGGTGAGGGTCATGGTGTG : 3990

Seqid3 : : 1950

4000 * 4020

Seqid1 : CTTTATGGTGAAATGGATTGGGCAGAAGAG : 4020

Seqid3 : : 1980

* 4040 *

Seqid1 : CTGTATGCTCGCCTTGGCGTATTTGGTGAA : 4050

26/45

Seqid3 : : 2010

4060 * 4080

Seqid1 : AATCTGACCGTCAGCCAGCCACCGCTTGTG : 4080

Seqid3 : : 2040

* 4100 *

Seqid1 : ACGGCACAAATCAGCCCTGAGCTGGAAGTG : 4110

Seqid3 : : 2070

4120 * 4140

Seqid1 : ATTATCAGACCTTTCCAAAAATTGTAGAT : 4140

Seqid3 : : 2100

* 4160 *

Seqid1 : GTTCAAGGCGTTGTTAGTATCCCATCAGCA : 4170

Seqid3 : : 2130

4180 * 4200

Seqid1 : ACCATTCGTCCGCCAGAGGCAACCGCAGAT : 4200

Seqid3 : : 2160

27/45

	*	4220	*
Seqid1	:	ATCGTAACCGAATCCCCAGATGTATCAGTC	: 4230
Seqid3	:	: 2190
		4240	*
Seqid1	:	ATTGACCGCCGTATCACAGGTAATATTGAC	: 4260
Seqid3	:	: 2220
	*	4280	*
Seqid1	:	CAGATTTTAAGTAGAGCGAAACCTTGGGAT	: 4290
Seqid3	:	: 2250
		4300	*
Seqid1	:	ATTAATGCAAATATCGGTGTTGATTGGGT	: 4320
Seqid3	:	: 2280
	*	4340	*
Seqid1	:	AATGAGATTGAATTTTCGTGGATTGGGGCG	: 4350
Seqid3	:	: 2310
		4360	*
			4380

28/45

Seqid1 : GTATTACCATTGGCAGGCGCGATACATTTG : 4380

Seqid3 : : 2340

* 4400 *

Seqid1 : ACTCAGTCAGGACAAGGTGCAATGCAAGCT : 4410

Seqid3 : : 2370

4420 * 4440

Seqid1 : CTTGGCGTGGTTCAGGTTTCTAAACGCACA : 4440

Seqid3 : : 2400

* 4460 *

Seqid1 : AAAATCGATGTCATCGGTCAAATTTGGAT : 4470

Seqid3 : : 2430

4480 * 4500

Seqid1 : TTAAATTATGCCCAAATTCGTTTTGATGGC : 4500

Seqid3 : : 2460

* 4520 *

Seqid1 : GACATGCTTAATCCACGCTTATCTATTGAG : 4530

Seqid3 : : 2490

29/45

	4540	*	4560
Seqid1	: GGTGAAAAACAAATTGAAGGGCAAACGGTG	:	4560
Seqid3	:	:	2520
	*	4580	*
Seqid1	: GGCGTTCGCATTCGTGGAACGGCATCAAAT	:	4590
Seqid3	:	:	2550
	4600	*	4620
Seqid1	: CCAAACATTACCGTATTTAATGATGCAGGT	:	4620
Seqid3	:	:	2580
	*	4640	*
Seqid1	: TTGGATGAATACCAAGCAATGAATGCACTG	:	4650
Seqid3	:	:	2610
	4660	*	4680
Seqid1	: GTGACAGGACGCATTAGCGAATCAAGTGAT	:	4680
Seqid3	:	:	2640

30/45

	*	4700	*
Seqid1	:	TTGGGTATCACAGAGCAGGGTTTTTCGCTCA	: 4710
Seqid3	:	: 2670
	4720	*	4740
Seqid1	:	CAAGTAACCAATCACCTTGCTGCTGCTGGC	: 4740
Seqid3	:	: 2700
	*	4760	*
Seqid1	:	TTAAGCTTAGGTTTGTTCAGGAACCAGAGAT	: 4770
Seqid3	:	: 2730
	4780	*	4800
Seqid1	:	TTAACCAACCAAATTGGTCACGCATTTGGT	: 4800
Seqid3	:	: 2760
	*	4820	*
Seqid1	:	TTTCAGAGTTTGACCATTGATGCCTCGGGC	: 4830
Seqid3	:	: 2790
	4840	*	4860
Seqid1	:	AGCTCAGATGATACCAATGTTAATGTTACA	: 4860

31/45

Seqid3 : : 2820

* 4880 *

Seqid1 : GGCTATATTACACCAGATTTGTATTTGCGT : 4890

Seqid3 : : 2850

4900 * 4920

Seqid1 : TATGGCGTGGGTGTGTTTAATGCTGAATCA : 4920

Seqid3 : : 2880

* 4940 *

Seqid1 : ACCTTATCCATGCGTTATCAATTGACACGC : 4950

Seqid3 : : 2910

4960 * 4980

Seqid1 : CGTGTTTATATTGAAGCAACCAGTGCCGCT : 4980

Seqid3 : : 2940

* 5000 *

Seqid1 : GAAAATATGGTTGATGTGATTTATCGTTGG : 5010

Seqid3 : : 2970

32/45

Seqid1	:	AAGTTTTAG	:	5019
Seqid3	:	:	2979

33/45

Figure 2 : Alignment of the BASB132 polypeptide sequences.

Identity to SeqID No:2 is indicated by a dot. Gap is indicated by a dash.

	*	20	*
Seqid2	:	MTHQDNSKNSSSAENGVSAGVAAPT	30
Seqid4	:	-----	-
	40	*	60
Seqid2	:	PPNPPKPQWLIYLIRLIVGAI	60
Seqid4	:	-----	-
	*	80	*
Seqid2	:	LFAMTNSESGSKFLIEKIALETG	90
Seqid4	:	-----	-
	100	*	120
Seqid2	:	GSIRHGVVVDVKIAQSE	120
Seqid4	:	-----	-
	*	140	*
Seqid2	:	LGWRALFARQVHLVNP	150
Seqid4	:	-----	-
	160	*	180

34/45

Seqid2 : GEPFDYATINLPVTLKLENAKVNEIIYDQV : 180

Seqid4 : ----- :

* 200 *

Seqid2 : DSEPVLHHIAFDHASWADSTVKIDNAMLS : 210

Seqid4 : ----- : -

220 * 240

Seqid2 : YGDDINISHATGGIDLTGHYPLSLADVHI : 240

Seqid4 : ----- : -

* 260 *

Seqid2 : LALDDAYFDTLSVKAGGSLKRTVGTTLTGKY : 270

Seqid4 : ----- : -

280 * 300

Seqid2 : NQHHVTGSFIAQGLDKNSPFSARLDFDEVR : 300

Seqid4 : ----- : -

* 320 *

Seqid2 : LPYADSQNILLKNGSIIADGVISNIELRIN : 330

Seqid4 : ----- : -

35/45

	340	*	360
Seqid2	: TELSAKDIPDGHYHGRGIIRGSTMQIPYLQ	:	360
Seqid4	: -----	:	-
	*	380	*
Seqid2	: ADTANGTLVATGDMTWEDGYELDATITADG	:	390
Seqid4	: -----	:	-
	400	*	420
Seqid2	: YRIREEMPSDYHEYRAYLPKVLTGSLGVKY	:	420
Seqid4	: -----	:	-
	*	440	*
Seqid2	: TLLDKASHDTRFEFDLNQKDGERIQATLAQ	:	450
Seqid4	: -----	:	-
	460	*	480
Seqid2	: NQQSDHEPWRIDATWANLIRHDIPQIGEIH	:	480
Seqid4	: -----	:	-
	*	500	*
Seqid2	: SRSGQASVRLENGHTYINASADIVKLNAV	:	510

36/45

Seqid4 : ----- : -

520 * 540

Seqid2 : SGSYHVQANIEQNQHLLHLLTDFNYQGVMGEL : 540

Seqid4 : ----- : -

* 560 *

Seqid2 : TGTGRVDFATAQRPLEWQLDVTANPVKPNA : 570

Seqid4 : ----- : -

580 * 600

Seqid2 : YFQTPNQTPFEQISGRRIASGRLREIDGVS : 600

Seqid4 : ----- : -

* 620 *

Seqid2 : IHDIKVDDSDLTALLNDGKQVHLMGQGVSK : 630

Seqid4 : ----- : -

640 * 660

Seqid2 : IRLQDGQISHLKANFDGKLAQDILPQVADS : 660

Seqid4 : ----- : -

37/45

	*	680	*
Seqid2	:	SIGLDIEGDLNDLTITRAVMMNDSGKVS LA	: 690
Seqid4	:	-----	: 10
		700	*
Seqid2	:	GRLGLNDGIHWQIQGRLDEVDTAA FVDDDN	: 720
Seqid4	:	: 40
	*	740	*
Seqid2	:	LVAIITGDLATSGRYRDGVMDDIALTFDGQ	: 750
Seqid4	:	: 70
		760	*
Seqid2	:	VINNSIPNGHLSIDATGSGNRFMVNHLRHD	: 780
Seqid4	:	: 100
	*	800	*
Seqid2	:	GTAGTLNATGWVDISQGA AWQLEADMNSLN	: 810
Seqid4	:	: 130
		820	*
			840

38/45

Seqid2 : LGAFIQSLDSDLNGTIQLAGNWQESHQVID : 840

Seqid4 : : 160

*

860

*

Seqid2 : ISNLDITGSYNNQPLIATGSLYAKLSLPKD : 870

Seqid4 : : 190

880

*

900

Seqid2 : LAGYIQSIKQASRPPTSSDDLALKNRDN : 900

Seqid4 : : 220

*

920

*

Seqid2 : NARQTQNIVHKLNADNLQVRIGNNHLAISG : 930

Seqid4 : : 250

940

*

960

Seqid2 : DERQLTASVNVIDLGQLIDTASGAIQGGVI : 960

Seqid4 : : 280

*

980

*

Seqid2 : VVNDHHALPTLYIDASVSSLRFANITIQNA : 990

39/45

Seqid4 : : 310

1000 * 1020

Seqid2 : QAIGKIVNLGNSESQLLVQGDDIIVMGRTI : 1020

Seqid4 : : 340

* 1040 *

Seqid2 : KSARMDFSGTEADHILSISTKSGDIEASMH : 1050

Seqid4 : : 370

1060 * 1080

Seqid2 : IDGAFNRNNMRYHGVLSDG FVKSGFGKMSQ : 1080

Seqid4 : : 400

* 1100 *

Seqid2 : RQPTEFSYGLDDNSLQIAAHCWQSAHIQGD : 1110

Seqid4 : : 430

1120 * 1140

Seqid2 : GVGAICLQDTLSYTPQSGNVNLI IQNLDTQ : 1140

Seqid4 : : 460

40/45

	*	1160	*
Seqid2	:	VLLAALPSDIHWKSMNLNGRIKAMWQAGQSP	: 1170
Seqid4	:	: 490
		1180	*
Seqid2	:	LVDAVLYSDDGTIGLTQEETGYIEMPYQRA	: 1200
Seqid4	:	: 520
	*	1220	*
Seqid2	:	SVIAKSVDNGLKVRTDILGTAGRGYADVII	: 1230
Seqid4	:	: 550
		1240	*
Seqid2	:	NPKQTDKPISGALVMNDLNLAVLRPFFPSI	: 1260
Seqid4	:	: 580
	*	1280	*
Seqid2	:	QTLSGKVSLAGGLGGTLSKPLFYGNANLKN	: 1290
Seqid4	:	: 610
		1300	*
			1320

Seqid2 : GALAIVGVMPISPVDATVNIRGTKARLDG : 1320

Seqid4 : : 640

* 1340 *

Seqid2 : RFMGGEHGVLYGEMDWAEELYARLGVFGE : 1350

Seqid4 : : 670

1360 * 1380

Seqid2 : NLTVSQPPLVTAQISPELEVIIRPFQKFVD : 1380

Seqid4 : : 700

* 1400 *

Seqid2 : VQGVVSIPSATIRPPEATADIVTESPDVSV : 1410

Seqid4 : : 730

1420 * 1440

Seqid2 : IDRRITGNIDQILSRAKPWDINANIGVDLG : 1440

Seqid4 : : 760

* 1460 *

Seqid2 : NEIEFRGFGAVLPLAGAIHLTQSGQGAMQA : 1470

Seqid4 : : 790

	1480	*	1500
Seqid2	: LGVVQVSKRTKIDVIGQNLDLNYAQIRFDG	:	1500
Seqid4	:	:	820
	*	1520	*
Seqid2	: DMLNPRLSIEGEKQIEGQTVGVRIRGTASN	:	1530
Seqid4	:	:	850
	1540	*	1560
Seqid2	: PNITVFNDAGLDEYQAMNALVTGRISSESD	:	1560
Seqid4	:	:	880
	*	1580	*
Seqid2	: LGITEQGFRSQVTNHLAAAGLSLGLSGTRD	:	1590
Seqid4	:	:	910
	1600	*	1620
Seqid2	: LTNQIGHAFGFQSLTIDASGSSDDTNVNVT	:	1620
Seqid4	:	:	940

```

*                               1640                               *
Seqid2  :  GYITPDLYLRYGVGVFNAESTLSMRYQLTR  :  1650
Seqid4  :  .....                               :  970

```

```

                                1660                                *
Seqid2      :      RYVIEATSAAENMVDVIYRWKF      :      1672
Seqid4      :      .....      :      992

```

44/45

Fig. 3A Small scale expression analysis of BASB132 :
Coomassie stained SDS- polyacrylamide gel.

Fig. 3B Small scale expression analysis of BASB132 :
Western-blotting (anti-His antibody).

Fig. 3A

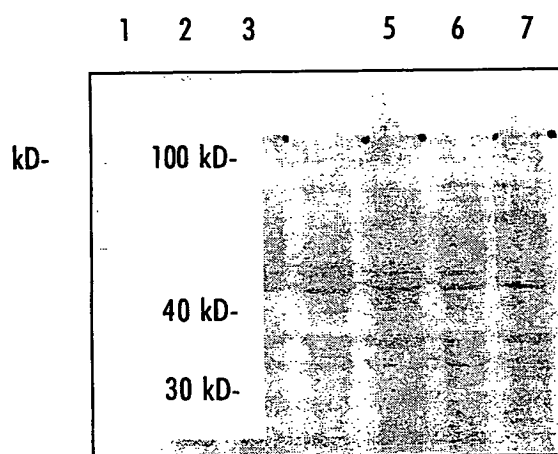
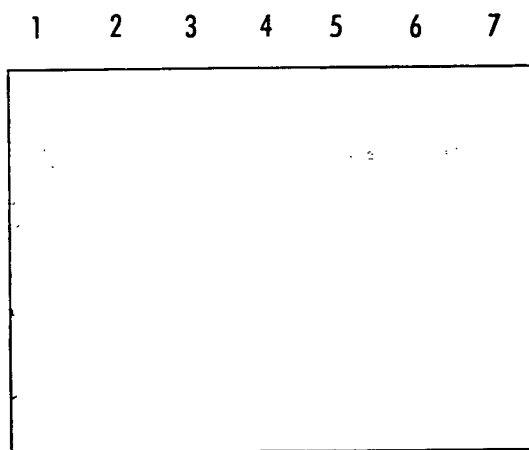


Fig. 3B



- 1: MW (kD)
- 2: clone 14, t=0h (no induction)
- 3: clone 14, t=4h (after induction)
- 4: clone 17, t=0h
- 5: clone 17, t=4h
- 6: clone 21, t=0h
- 7: clone 21, t=4h

45/45

Fig. 4A Coomassie stained SDS-polyacrylamide gel of purified BASB132

Fig. 4B Western-blotting of purified BASB132 (anti-His antibody).

Fig. 4A

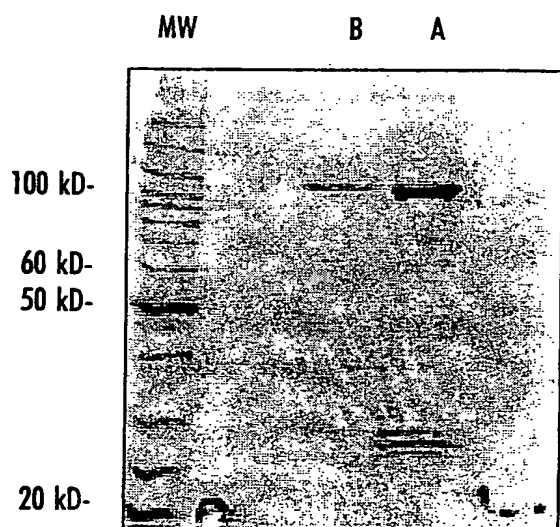
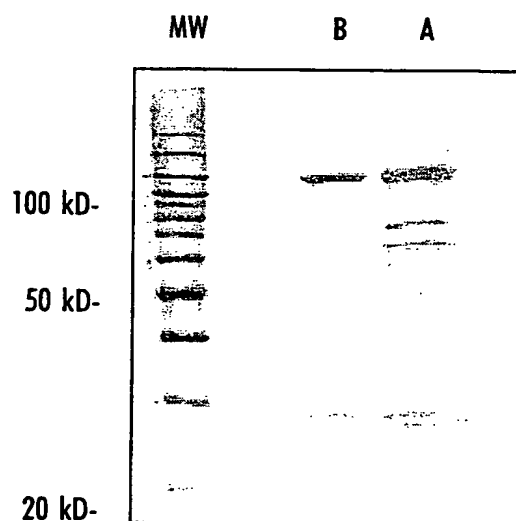


Fig. 4B



SEQUENCE LISTING

<110> SmithKline Beecham Biologicals S.A.

<120> Novel Compounds

<130> BM45417

<160> 13

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 5019

<212> DNA

<213> *Moraxella catarrhalis*

<400> 1

atgactcatc	aggataattc	aaaaaactcg	tcatcggcag	aaaatggggt	ttctgctggc	60
gttgctgcac	cgactaaaaac	tcagcccaaa	ccaccaaaacc	cacccaaacc	acaatggcta	120
atctatttaa	ttcgcttaat	tgctcgagca	attattgccg	ctttattgtt	gttggcggtg	180
ttatttgcca	tgaccaacag	cgagtcagggt	tcaaagtttt	taatcgaaaa	aattgctgta	240
gaaactggca	ctaagcttaa	gtacagcgaa	ggctcaattc	gccatggagt	ttgggtgcaa	300
gatgtcaaaa	tcgctcaaag	tgaagatatt	accatcacca	ttaatcgtgc	ctatgtacag	360
cttgggtggc	gagccttggt	tgctcgccaa	gtgcatttgg	tcaatcctaa	gattgataaa	420
gtttatgtga	caaacaccaa	gccatcaaca	ggcgaaccct	ttgattatgc	gaccatcaac	480
ctaccagtga	cgcttaagct	tgaaaatgcc	aaagtcaatg	aaattatcta	tgaccaagtg	540
gattctgagc	ctgtcgtact	gcatcatatc	gcatttgatc	acgcatcatg	ggcagattca	600
acagttaaaa	ttgataacgc	catgctaagc	tatggtgatg	atattaatat	cagccatgcc	660
actggtggaa	ttgatttaac	aggtcattat	ccgctgtcgt	tgctcggcaga	tgtgcatatt	720
ttggcacttg	atgatgcgta	ttttgatact	ttgtcgggtga	aagcaggcgg	tagccttaag	780
cgtaccggtg	gtacactcac	tggcaaatat	aatcagcatc	atgtgacagg	cagctttatc	840
gctcaagggt	tggataaaaa	ctcacctttt	agcgcacgcc	ttgattttga	tgaagtacga	900
ttgccttatg	ctgacagtca	aaatatttta	ctaaaaaatg	gctctatcat	cgctgatggc	960
gtcatctcaa	atatcgagct	acgcattaac	actgagttat	ccgccaaaaga	tattcctgat	1020
gggcattatc	acggtcgtgg	aattattcgt	ggcagtagca	tgcaaatccc	atatttgtag	1080
gctgatactg	ccaatggtag	tttgggtggc	acgggtgata	tgacttggga	agatggctat	1140
gagcttgatg	ccaccattac	agcagacggc	tatcgtatcc	gtgaagagat	gccaagtgat	1200
tatcatgaat	atagagccta	tctacctaag	gttttgacag	gctcacttgg	ggttaagtat	1260
acgctattag	acaaagccag	tcattgatact	cggtttgagt	ttgatctgaa	tcaaaaagac	1320
gggtgaacgca	ttcaagcgac	gctggctcaa	aatcaacaga	gtgatcatga	gccttggcgt	1380
attgatgcga	cttggggcaaa	tctaaccgcg	catgatattc	cacaaattgg	cgagattcat	1440
agccgctcag	gtcaggcatc	agttcgtttg	gagaacggac	atacctatat	taacgcctct	1500
gctgacattg	ttaaacttaa	tgccgtccca	agcggatcat	atcatgtgca	agcaaatatt	1560
gagcaaaatc	agcacttgca	cttgacagat	tttaactatc	aagggtgtgat	gggcgagttg	1620
acaggtagctg	gcagggttga	ttttgcgact	gcccagaagac	cgcttgagtg	gcagcttgat	1680
gtcactgcca	atccagtcaa	acccaatgcg	tattttcaaa	cacccaatca	aacaccgttt	1740
gagcaaaattt	caggcagaat	aattgcttca	gggcgtctgc	gtgagataga	tggtgtgtca	1800
attcatgaca	ttaaagtggg	tgacagtgat	ttaaccgctt	tattaaatga	cggcaagcaa	1860
gtacatctga	tgggtcaagg	cgtagcgaag	attaggctac	aagatggaca	aattagccat	1920
cttaaggcaa	attttgatgg	caagcttgct	caagatatct	taccgcaggt	agctgactca	1980
agcattggat	tggatattga	aggtgatttg	aatgatttga	ccatcacacg	ggcagtcagt	2040
atgaatgact	caggcaaaagt	atcgcttgca	ggctcgtctg	gattaaatga	cgggtattcat	2100
tggcaaaattc	aaggggcgatt	agatgaggtg	gataccgccg	cttttggtga	tgatgacaat	2160
ttagttgcca	tcattcacag	cgatcttgcc	acatcaggca	gatatcgtga	tggcgtgatg	2220
gatgatatcg	cactcacttt	tgatgggtcaa	gtgattaaca	attcaatccc	aaatgggtcat	2280
ctaagtattg	atgtagcagg	ttctggcaat	cgctttatgg	tcaatcatct	aagacatgat	2340

```

ggcacagcag gcacgctgaa tgcgacaggg tgggtagata ttagccaagg tgctgcatgg 2400
cagctagagg ctgatatgaa ttcatatgaat ttgggtgcat ttatccagtc gctagatact 2460
gacttaaacg gtacaattca gcttgacagg aactggcaag agtctcatca agtcattgat 2520
atcagtaatt tggatattac aggtagctac aataatcagc cacttatcgc cacaggcagc 2580
ttatatgcca agctttccct acctaaagat ttggctgggt atatccaaag tatcaagcaa 2640
gcaagccgcc caccgacttc gtctgatgat ttattggcac taaaaaatcg tattgataac 2700
aacgcacgcc aaacccaaaa tatcgttcat aaactcaatg ctgataatct acaagtccgc 2760
attggcaata atcatttggc catatcaggg gatgaaaagac agttgaccgc cagtgtgaat 2820
gtgattgatc ttgggcagct gattgatacc gcaagtgggt cgattcaagg cgggtgtgatt 2880
gtcgtcaatg atcatcacgc tttgccaaact ttatatattg atgccagtgt ctcatcactc 2940
agatttgcca atatcaccat ccaaaacgcc caagctatcg gtaaaattgt caatctgggt 3000
aatagcgaaa gccagctggt ggttcagggt gatgatatta tcgtgatggg acgcaccatt 3060
aaatctgctc ggatggattt tagtggcaca gaagccgatc atatcttattc aatttcaacc 3120
aaaagcgggt atatcgaagc atctatgcat attgatgggg catttaatcg caacaatatg 3180
cgatatcaat gtgttttgc tgatgggttt tcaagagtg gatttggcaa aatgtcacag 3240
cgtcagccga ccgagtttag ctatggattg gatgataata gcttacaat tgcagcgcat 3300
tgttggcaat cggcacatat ccaaggcgat ggctgggtg cgatttgttt acaagataca 3360
ttaagttata caccgcaatc aggggaatgtc aatcttatta tccaaaatct tgataactcaa 3420
gtactatttg cggctctacc cagtgatatt cattggaaat ccatgctcaa tggctgcatt 3480
aaggcaatgt ggcaggcagg tcaatcacct ttggttgatg ctgtactgta ttctgatgat 3540
ggtaacaattg gcttaactca agaagagaca gtttatattg agatgccata tcagcgtgag 3600
tctgtgattg ccaaaagtgt tgataatggc ttaaaggtaa gaaccgatat cttgggtacg 3660
gcaggtcgtg gttatgccga tgtaatcacc aacccaagc aaacagacaa gccattttct 3720
gggtgcgctt ttatgaatga tcttaatttg gcggtattgc gaccattttt cccgagcatt 3780
caaacattga gcggtaaagt cagcttggca ggtggcttag gtggtacatt atccaaacca 3840
ttgttttatg gtaatgccaa tttaaagaat ggtgcattag cgattgttgg cgtaccaatg 3900
cccatttctg atgttgatgc caggtgaac atacgaggtg cttaaagcag cttagacggg 3960
agatttatgg gtgtgaggg tcatgggtg ctttatgggt aaatggattg ggcagaagag 4020
ctgtatgctc gccttggcgt atttggtgaa aatctgaccg tcagccagcc accgcttgtg 4080
acggcacaaa tcagccctga gctggaagtg attatcagac ctttccaaa atttgtagat 4140
gttcaaggcg ttgttagtat cccatcagca accattcgtc cgccagaggc aaccgcagat 4200
atcgtaaccg aatccccaga tgtatcagtc attgaccgcc gtatcacagg taatattgac 4260
cagattttaa gtagagcgaa accttgggat attaatgcaa atatcggtgt tgatttgggt 4320
aatgagattg aatttcgtgg atttggggcg gtattaccat tggcaggcgc gatcatttg 4380
actcagtcag gacaagggtg aatgcaagct cttggcgtgg ttcagggttc taaacgcaca 4440
aaaatcgatg tcatcggtca aaatttggat ttaaattatg cccaaattcg ttttgatggc 4500
gacatgctta atccacgctt atctattgag ggtgaaaaac aaattgaagg gcaaacgggt 4560
ggcgttcgca ttcgtggaac ggcatacaat ccaaacatta ccgtatttaa tgatgcaggt 4620
ttggatgaat accaagcaat gaatgcactg gtgacaggac gcattagcga atcaagtgat 4680
ttgggtatca cagagcaggg ttttcgctca caagtaacca atcaccttgc tgctgtggc 4740
ttaagcttag gtttgtcagg aaccagagat ttaaccaacc aaattgggtc cgcatttggg 4800
tttcagagtt tgaccattga tgcctcgggc agctcagatg ataccaatgt taatgttaca 4860
ggctatatta caccagattt gtatttgcgt tatggcgtgg gtgtgtttaa tgctgaatca 4920
accttatcca tgcgttatca attgacacgc cgtgtttata ttgaagcaac cagtgccgct 4980
gaaaaatatg ttgatgtgat ttatcgttgg aagtttttag 5019

```

<210> 2

<211> 1672

<212> PRT

<213> *Moraxella catarrhalis*

<400> 2

```

Met Thr His Gln Asp Asn Ser Lys Asn Ser Ser Ser Ala Glu Asn Gly
 1           5           10          15
Val Ser Ala Gly Val Ala Ala Pro Thr Lys Thr Gln Pro Lys Pro Pro
      20           25           30

```

Asn Pro Pro Lys Pro Gln Trp Leu Ile Tyr Leu Ile Arg Leu Ile Val
 35 40 45
 Gly Ala Ile Ile Ala Ala Leu Leu Leu Leu Ala Val Leu Phe Ala Met
 50 55 60
 Thr Asn Ser Glu Ser Gly Ser Lys Phe Leu Ile Glu Lys Ile Ala Leu
 65 70 75 80
 Glu Thr Gly Thr Lys Leu Lys Tyr Ser Glu Gly Ser Ile Arg His Gly
 85 90 95
 Val Trp Val Gln Asp Val Lys Ile Ala Gln Ser Glu Asp Ile Thr Ile
 100 105 110
 Thr Ile Asn Arg Ala Tyr Val Gln Leu Gly Trp Arg Ala Leu Phe Ala
 115 120 125
 Arg Gln Val His Leu Val Asn Pro Lys Ile Asp Lys Val Tyr Val Thr
 130 135 140
 Asn Thr Lys Pro Ser Thr Gly Glu Pro Phe Asp Tyr Ala Thr Ile Asn
 145 150 155 160
 Leu Pro Val Thr Leu Lys Leu Glu Asn Ala Lys Val Asn Glu Ile Ile
 165 170 175
 Tyr Asp Gln Val Asp Ser Glu Pro Val Val Leu His His Ile Ala Phe
 180 185 190
 Asp His Ala Ser Trp Ala Asp Ser Thr Val Lys Ile Asp Asn Ala Met
 195 200 205
 Leu Ser Tyr Gly Asp Asp Ile Asn Ile Ser His Ala Thr Gly Gly Ile
 210 215 220
 Asp Leu Thr Gly His Tyr Pro Leu Ser Leu Ser Ala Asp Val His Ile
 225 230 235 240
 Leu Ala Leu Asp Asp Ala Tyr Phe Asp Thr Leu Ser Val Lys Ala Gly
 245 250 255
 Gly Ser Leu Lys Arg Thr Val Gly Thr Leu Thr Gly Lys Tyr Asn Gln
 260 265 270
 His His Val Thr Gly Ser Phe Ile Ala Gln Gly Leu Asp Lys Asn Ser
 275 280 285
 Pro Phe Ser Ala Arg Leu Asp Phe Asp Glu Val Arg Leu Pro Tyr Ala
 290 295 300
 Asp Ser Gln Asn Ile Leu Leu Lys Asn Gly Ser Ile Ile Ala Asp Gly
 305 310 315 320
 Val Ile Ser Asn Ile Glu Leu Arg Ile Asn Thr Glu Leu Ser Ala Lys
 325 330 335
 Asp Ile Pro Asp Gly His Tyr His Gly Arg Gly Ile Ile Arg Gly Ser
 340 345 350
 Thr Met Gln Ile Pro Tyr Leu Gln Ala Asp Thr Ala Asn Gly Thr Leu
 355 360 365
 Val Ala Thr Gly Asp Met Thr Trp Glu Asp Gly Tyr Glu Leu Asp Ala
 370 375 380
 Thr Ile Thr Ala Asp Gly Tyr Arg Ile Arg Glu Glu Met Pro Ser Asp
 385 390 395 400
 Tyr His Glu Tyr Arg Ala Tyr Leu Pro Lys Val Leu Thr Gly Ser Leu
 405 410 415
 Gly Val Lys Tyr Thr Leu Leu Asp Lys Ala Ser His Asp Thr Arg Phe
 420 425 430
 Glu Phe Asp Leu Asn Gln Lys Asp Gly Glu Arg Ile Gln Ala Thr Leu
 435 440 445
 Ala Gln Asn Gln Gln Ser Asp His Glu Pro Trp Arg Ile Asp Ala Thr
 450 455 460
 Trp Ala Asn Leu Ile Arg His Asp Ile Pro Gln Ile Gly Glu Ile His
 465 470 475 480
 Ser Arg Ser Gly Gln Ala Ser Val Arg Leu Glu Asn Gly His Thr Tyr
 485 490 495

```

Ile Asn Ala Ser Ala Asp Ile Val Lys Leu Asn Ala Val Pro Ser Gly
500 505 510
Ser Tyr His Val Gln Ala Asn Ile Glu Gln Asn Gln His Leu His Leu
515 520 525
Thr Asp Phe Asn Tyr Gln Gly Val Met Gly Glu Leu Thr Gly Thr Gly
530 535 540
Arg Val Asp Phe Ala Thr Ala Gln Arg Pro Leu Glu Trp Gln Leu Asp
545 550 555 560
Val Thr Ala Asn Pro Val Lys Pro Asn Ala Tyr Phe Gln Thr Pro Asn
565 570 575
Gln Thr Pro Phe Glu Gln Ile Ser Gly Arg Ile Ile Ala Ser Gly Arg
580 585 590
Leu Arg Glu Ile Asp Gly Val Ser Ile His Asp Ile Lys Val Asp Asp
595 600 605
Ser Asp Leu Thr Ala Leu Leu Asn Asp Gly Lys Gln Val His Leu Met
610 615 620
Gly Gln Gly Val Ser Lys Ile Arg Leu Gln Asp Gly Gln Ile Ser His
625 630 635 640
Leu Lys Ala Asn Phe Asp Gly Lys Leu Ala Gln Asp Ile Leu Pro Gln
645 650 655
Val Ala Asp Ser Ser Ile Gly Leu Asp Ile Glu Gly Asp Leu Asn Asp
660 665 670
Leu Thr Ile Thr Arg Ala Val Met Met Asn Asp Ser Gly Lys Val Ser
675 680 685
Leu Ala Gly Arg Leu Gly Leu Asn Asp Gly Ile His Trp Gln Ile Gln
690 695 700
Gly Arg Leu Asp Glu Val Asp Thr Ala Ala Phe Val Asp Asp Asp Asn
705 710 715 720
Leu Val Ala Ile Ile Thr Gly Asp Leu Ala Thr Ser Gly Arg Tyr Arg
725 730 735
Asp Gly Val Met Asp Asp Ile Ala Leu Thr Phe Asp Gly Gln Val Ile
740 745 750
Asn Asn Ser Ile Pro Asn Gly His Leu Ser Ile Asp Ala Thr Gly Ser
755 760 765
Gly Asn Arg Phe Met Val Asn His Leu Arg His Asp Gly Thr Ala Gly
770 775 780
Thr Leu Asn Ala Thr Gly Trp Val Asp Ile Ser Gln Gly Ala Ala Trp
785 790 795 800
Gln Leu Glu Ala Asp Met Asn Ser Leu Asn Leu Gly Ala Phe Ile Gln
805 810 815
Ser Leu Asp Thr Asp Leu Asn Gly Thr Ile Gln Leu Ala Gly Asn Trp
820 825 830
Gln Glu Ser His Gln Val Ile Asp Ile Ser Asn Leu Asp Ile Thr Gly
835 840 845
Ser Tyr Asn Asn Gln Pro Leu Ile Ala Thr Gly Ser Leu Tyr Ala Lys
850 855 860
Leu Ser Leu Pro Lys Asp Leu Ala Gly Tyr Ile Gln Ser Ile Lys Gln
865 870 875 880
Ala Ser Arg Pro Pro Thr Ser Ser Asp Asp Leu Leu Ala Leu Lys Asn
885 890 895
Arg Ile Asp Asn Asn Ala Arg Gln Thr Gln Asn Ile Val His Lys Leu
900 905 910
Asn Ala Asp Asn Leu Gln Val Arg Ile Gly Asn Asn His Leu Ala Ile
915 920 925
Ser Gly Asp Glu Arg Gln Leu Thr Ala Ser Val Asn Val Ile Asp Leu
930 935 940
Gly Gln Leu Ile Asp Thr Ala Ser Gly Ala Ile Gln Gly Gly Val Ile
945 950 955 960

```

Val	Val	Asn	Asp	His	His	Ala	Leu	Pro	Thr	Leu	Tyr	Ile	Asp	Ala	Ser	965	970	975
Val	Ser	Ser	Leu	Arg	Phe	Ala	Asn	Ile	Thr	Ile	Gln	Asn	Ala	Gln	Ala	980	985	990
Ile	Gly	Lys	Ile	Val	Asn	Leu	Gly	Asn	Ser	Glu	Ser	Gln	Leu	Leu	Val	995	1000	1005
Gln	Gly	Asp	Asp	Ile	Ile	Val	Met	Gly	Arg	Thr	Ile	Lys	Ser	Ala	Arg	1010	1015	1020
Met	Asp	Phe	Ser	Gly	Thr	Glu	Ala	Asp	His	Ile	Leu	Ser	Ile	Ser	Thr	1025	1030	1035
Lys	Ser	Gly	Asp	Ile	Glu	Ala	Ser	Met	His	Ile	Asp	Gly	Ala	Phe	Asn	1045	1050	1055
Arg	Asn	Asn	Met	Arg	Tyr	His	Gly	Val	Leu	Ser	Asp	Gly	Phe	Val	Lys	1060	1065	1070
Ser	Gly	Phe	Gly	Lys	Met	Ser	Gln	Arg	Gln	Pro	Thr	Glu	Phe	Ser	Tyr	1075	1080	1085
Gly	Leu	Asp	Asp	Asn	Ser	Leu	Gln	Ile	Ala	Ala	His	Cys	Trp	Gln	Ser	1090	1095	1100
Ala	His	Ile	Gln	Gly	Asp	Gly	Val	Gly	Ala	Ile	Cys	Leu	Gln	Asp	Thr	1105	1110	1115
Leu	Ser	Tyr	Thr	Pro	Gln	Ser	Gly	Asn	Val	Asn	Leu	Ile	Ile	Gln	Asn	1125	1130	1135
Leu	Asp	Thr	Gln	Val	Leu	Leu	Ala	Ala	Leu	Pro	Ser	Asp	Ile	His	Trp	1140	1145	1150
Lys	Ser	Met	Leu	Asn	Gly	Arg	Ile	Lys	Ala	Met	Trp	Gln	Ala	Gly	Gln	1155	1160	1165
Ser	Pro	Leu	Val	Asp	Ala	Val	Leu	Tyr	Ser	Asp	Asp	Gly	Thr	Ile	Gly	1170	1175	1180
Leu	Thr	Gln	Glu	Glu	Thr	Gly	Tyr	Ile	Glu	Met	Pro	Tyr	Gln	Arg	Ala	1185	1190	1195
Ser	Val	Ile	Ala	Lys	Ser	Val	Asp	Asn	Gly	Leu	Lys	Val	Arg	Thr	Asp	1205	1210	1215
Ile	Leu	Gly	Thr	Ala	Gly	Arg	Gly	Tyr	Ala	Asp	Val	Ile	Ile	Asn	Pro	1220	1225	1230
Lys	Gln	Thr	Asp	Lys	Pro	Ile	Ser	Gly	Ala	Leu	Val	Met	Asn	Asp	Leu	1235	1240	1245
Asn	Leu	Ala	Val	Leu	Arg	Pro	Phe	Phe	Pro	Ser	Ile	Gln	Thr	Leu	Ser	1250	1255	1260
Gly	Lys	Val	Ser	Leu	Ala	Gly	Gly	Leu	Gly	Gly	Thr	Leu	Ser	Lys	Pro	1265	1270	1275
Leu	Phe	Tyr	Gly	Asn	Ala	Asn	Leu	Lys	Asn	Gly	Ala	Leu	Ala	Ile	Val	1285	1290	1295
Gly	Val	Pro	Met	Pro	Ile	Ser	Asp	Val	Asp	Ala	Thr	Val	Asn	Ile	Arg	1300	1305	1310
Gly	Thr	Lys	Ala	Arg	Leu	Asp	Gly	Arg	Phe	Met	Gly	Gly	Glu	Gly	His	1315	1320	1325
Gly	Val	Leu	Tyr	Gly	Glu	Met	Asp	Trp	Ala	Glu	Glu	Leu	Tyr	Ala	Arg	1330	1335	1340
Leu	Gly	Val	Phe	Gly	Glu	Asn	Leu	Thr	Val	Ser	Gln	Pro	Pro	Leu	Val	1345	1350	1355
Thr	Ala	Gln	Ile	Ser	Pro	Glu	Leu	Glu	Val	Ile	Arg	Pro	Phe	Gln		1365	1370	1375
Lys	Phe	Val	Asp	Val	Gln	Gly	Val	Val	Ser	Ile	Pro	Ser	Ala	Thr	Ile	1380	1385	1390
Arg	Pro	Pro	Glu	Ala	Thr	Ala	Asp	Ile	Val	Thr	Glu	Ser	Pro	Asp	Val	1395	1400	1405
Ser	Val	Ile	Asp	Arg	Arg	Ile	Thr	Gly	Asn	Ile	Asp	Gln	Ile	Leu	Ser	1410	1415	1420

Arg Ala Lys Pro Trp Asp Ile Asn Ala Asn Ile Gly Val Asp Leu Gly
 1425 1430 1435 144
 Asn Glu Ile Glu Phe Arg Gly Phe Gly Ala Val Leu Pro Leu Ala Gly
 1445 1450 1455
 Ala Ile His Leu Thr Gln Ser Gly Gln Gly Ala Met Gln Ala Leu Gly
 1460 1465 1470
 Val Val Gln Val Ser Lys Arg Thr Lys Ile Asp Val Ile Gly Gln Asn
 1475 1480 1485
 Leu Asp Leu Asn Tyr Ala Gln Ile Arg Phe Asp Gly Asp Met Leu Asn
 1490 1495 1500
 Pro Arg Leu Ser Ile Glu Gly Glu Lys Gln Ile Glu Gly Gln Thr Val
 1505 1510 1515 152
 Gly Val Arg Ile Arg Gly Thr Ala Ser Asn Pro Asn Ile Thr Val Phe
 1525 1530 1535
 Asn Asp Ala Gly Leu Asp Glu Tyr Gln Ala Met Asn Ala Leu Val Thr
 1540 1545 1550
 Gly Arg Ile Ser Glu Ser Ser Asp Leu Gly Ile Thr Glu Gln Gly Phe
 1555 1560 1565
 Arg Ser Gln Val Thr Asn His Leu Ala Ala Ala Gly Leu Ser Leu Gly
 1570 1575 1580
 Leu Ser Gly Thr Arg Asp Leu Thr Asn Gln Ile Gly His Ala Phe Gly
 1585 1590 1595 160
 Phe Gln Ser Leu Thr Ile Asp Ala Ser Gly Ser Ser Asp Asp Thr Asn
 1605 1610 1615
 Val Asn Val Thr Gly Tyr Ile Thr Pro Asp Leu Tyr Leu Arg Tyr Gly
 1620 1625 1630
 Val Gly Val Phe Asn Ala Glu Ser Thr Leu Ser Met Arg Tyr Gln Leu
 1635 1640 1645
 Thr Arg Arg Val Tyr Ile Glu Ala Thr Ser Ala Ala Glu Asn Met Val
 1650 1655 1660
 Asp Val Ile Tyr Arg Trp Lys Phe
 1665 1670

<210> 3

<211> 2979

<212> DNA

<213> *Moraxella catarrhalis*

<400> 3

atgaatgact	caggcaaagt	atcgcttgca	ggcgtctctg	gattaaatga	cggtattcat	60
tggcaaattc	aagggcgatt	agatgaggtg	gataccgccg	cttttggtga	tgatgacaat	120
ttagttgcga	tcatcacagg	cgatcttgcc	acatcaggca	gatatcgtga	tggcgtgatg	180
gatgatatcg	cactcacttt	tgatgggtcaa	gtgattaaca	attcaatccc	aaatgggtcat	240
ctaagtattg	atgcgacagg	ttctggcaat	cgctttatgg	tcaatcatct	aagacatgat	300
ggcacagcag	gcacgctgaa	tgcgacaggc	tgggtagata	ttagccaagg	tgctgcatgg	360
cagctagagg	ctgatatgaa	ttcattgaat	ttgggtgcat	ttatccagtc	gctagataact	420
gacttaaacg	gtacaattca	gcttgcaggc	aactggcaag	agtctcatca	agtcattgat	480
atcagtaatt	tggatattac	aggtagctac	aataatcagc	cacttatcgc	cacaggcagc	540
ttatatgccca	agctttccct	acctaaagat	ttggctgggt	atatccaaag	tatcaagcaa	600
gcaagccgcc	caccgacttc	gtctgatgat	ttattggcac	taaaaaatcg	tattgataac	660
aacgcacgcc	aaacccaaaa	tatcgttcat	aaactcaatg	ctgataatct	acaagtcgcg	720
attggcaata	atcatttggc	catatcaggc	gatgaaagac	agttgaccgc	cagtgtgaat	780
gtgattgatc	ttgggcagct	gattgatacc	gcaagtgggtg	cgattcaagg	cggtgtgatt	840
gtcgtcaatg	atcatcacgc	tttgccaact	ttatatattg	atgccagtgt	ctcatcactc	900
agatttgcca	atatcaccat	ccaaaacgcc	caagctatcg	gtaaaattgt	caatctgggt	960
aatagcgaaa	gccagctggt	ggttcagggt	gatgatatta	tcgtgatggg	acgcaccatt	1020
aaatctgctc	ggatggattt	tagtggcaca	gaagccgata	atatcttatc	aatttcaacc	1080
aaaagcgggtg	atatcgaagc	atctatgcat	attgatgggg	catttaatcg	caacaatatg	1140

```

cgatatcatg gtgttttgtc tgatggtttt gtcaagagtg gatttggcaa aatgtcacag 1200
cgtcagccga ccgagtttag ctatggattg gatgataata gcttacaaat tgcagcgcatt 1260
tggttggaat cggcacatat ccaaggcgat ggcgtgggtg cgatttgttt acaagataca 1320
ttaagttata caccgcaatc agggaatgtc aatcttatta tccaaaatct tgataactcaa 1380
gtactatttg cggctctacc cagtgatatt cattggaaat ccattgctcaa tgggtcgatt 1440
aaggcaatgt ggcaggcagg tcaatcacct ttggttgatg ctgtactgta ttctgatgat 1500
ggtacaattg gcttaactca agaagagaca gggttatatt agatgccata tcagcgtgcg 1560
tctgtgattg ccaaaagtgt tgataatggc ttaaaggtaa gaaccgatat cttgggtacg 1620
gcaggctcgtg gttatgccga tgtaatcatc aaccccaagc aaacagacaa gccattttct 1680
ggtgcgcttg ttatgaatga tcttaatttg gcggtattgc gaccattttt cccgagcatt 1740
caaacattga gcggtaaagt cagcttgcca ggtggcttag gtggtacatt atccaaacca 1800
ttgttttatg gtaatgccaa tttaaagaat ggtgcattag cgattgttgg cgtaccaatg 1860
cccatttctg atgttgatgc cacggtgaac atacgaggtg ctaaagcacg cttagacggg 1920
agattttatg gtggtgaggg tcatggtgtg ctttatgggt aaatggattg ggcagaagag 1980
ctgtatgctc gccttggcgt atttggtgaa aatctgaccg tcagccagcc accgcttggt 2040
acggcacaaa tcagccctga gctggaagtg attatcagac ctttccaaaa atttgtagat 2100
gttcaaggcg ttgttagtat cccatcagca accattcgtc cgccagaggc aaccgcagat 2160
atcgtaaccg aatccccaga tgtatcagtc attgaccgcc gtatcacagg taatattgac 2220
cagattttaa gtagagcgaa accttgggat attaatgcaa atatcgggtg tgatttgggt 2280
aatgagattg aatttcgtgg atttggggcg gtattaccat tggcaggcgc gatacatttg 2340
actcagtcag gacaagggtg aatgcaagct cttggcgtgg ttcaggtttc taaacgcaca 2400
aaaatcgatg tcatcggtca aaatttggat ttaaattatg cccaaattcg ttttgatggc 2460
gacatgctta atccacgctt atctattgag ggtgaaaaac aaattgaagg gcaaacggtg 2520
ggcgttcgca ttcgtggaac ggcacaaat ccaaacatta ccgtatttaa tgatgcaggt 2580
ttggatgaat accaagcaat gaatgcactg gtgacaggac gcattagcga atcaagtgat 2640
ttgggtatca cagagcaggg ttttcgctca caagtaacca atcaccttgc tgctgctggc 2700
ttaagcttag gtttgtcagg aaccagagat ttaaccaacc aaattggtca cgcatttggg 2760
tttcagagtt tgaccattga tgcctcgggc agctcagatg ataccaatgt taatgttaca 2820
ggctatatta caccagattt gtatttgcgt tatggcgtgg gtgtgtttaa tgctgaatca 2880
accttatcca tgcgttatca attgacacgc cgtgtttata ttgaagcaac cagtgccgct 2940
gaaaatatgg ttgatgtgat ttatcgttgg aagttttag 2979

```

<210> 4

<211> 992

<212> PRT

<213> *Moraxella catarrhalis*

<400> 4

```

Met Asn Asp Ser Gly Lys Val Ser Leu Ala Gly Arg Leu Gly Leu Asn
 1          5          10          15
Asp Gly Ile His Trp Gln Ile Gln Gly Arg Leu Asp Glu Val Asp Thr
          20          25          30
Ala Ala Phe Val Asp Asp Asp Asn Leu Val Ala Ile Ile Thr Gly Asp
          35          40          45
Leu Ala Thr Ser Gly Arg Tyr Arg Asp Gly Val Met Asp Asp Ile Ala
          50          55          60
Leu Thr Phe Asp Gly Gln Val Ile Asn Asn Ser Ile Pro Asn Gly His
          65          70          75          80
Leu Ser Ile Asp Ala Thr Gly Ser Gly Asn Arg Phe Met Val Asn His
          85          90          95
Leu Arg His Asp Gly Thr Ala Gly Thr Leu Asn Ala Thr Gly Trp Val
          100          105          110
Asp Ile Ser Gln Gly Ala Ala Trp Gln Leu Glu Ala Asp Met Asn Ser
          115          120          125
Leu Asn Leu Gly Ala Phe Ile Gln Ser Leu Asp Thr Asp Leu Asn Gly
          130          135          140
Thr Ile Gln Leu Ala Gly Asn Trp Gln Glu Ser His Gln Val Ile Asp
          145          150          155          160

```

Ile Ser Asn Leu Asp Ile Thr Gly Ser Tyr Asn Asn Gln Pro Leu Ile
 165 170 175
 Ala Thr Gly Ser Leu Tyr Ala Lys Leu Ser Leu Pro Lys Asp Leu Ala
 180 185 190
 Gly Tyr Ile Gln Ser Ile Lys Gln Ala Ser Arg Pro Pro Thr Ser Ser
 195 200 205
 Asp Asp Leu Leu Ala Leu Lys Asn Arg Ile Asp Asn Asn Ala Arg Gln
 210 215 220
 Thr Gln Asn Ile Val His Lys Leu Asn Ala Asp Asn Leu Gln Val Arg
 225 230 235 240
 Ile Gly Asn Asn His Leu Ala Ile Ser Gly Asp Glu Arg Gln Leu Thr
 245 250 255
 Ala Ser Val Asn Val Ile Asp Leu Gly Gln Leu Ile Asp Thr Ala Ser
 260 265 270
 Gly Ala Ile Gln Gly Gly Val Ile Val Val Asn Asp His His Ala Leu
 275 280 285
 Pro Thr Leu Tyr Ile Asp Ala Ser Val Ser Ser Leu Arg Phe Ala Asn
 290 295 300
 Ile Thr Ile Gln Asn Ala Gln Ala Ile Gly Lys Ile Val Asn Leu Gly
 305 310 315 320
 Asn Ser Glu Ser Gln Leu Leu Val Gln Gly Asp Asp Ile Ile Val Met
 325 330 335
 Gly Arg Thr Ile Lys Ser Ala Arg Met Asp Phe Ser Gly Thr Glu Ala
 340 345 350
 Asp His Ile Leu Ser Ile Ser Thr Lys Ser Gly Asp Ile Glu Ala Ser
 355 360 365
 Met His Ile Asp Gly Ala Phe Asn Arg Asn Asn Met Arg Tyr His Gly
 370 375 380
 Val Leu Ser Asp Gly Phe Val Lys Ser Gly Phe Gly Lys Met Ser Gln
 385 390 395 400
 Arg Gln Pro Thr Glu Phe Ser Tyr Gly Leu Asp Asp Asn Ser Leu Gln
 405 410 415
 Ile Ala Ala His Cys Trp Gln Ser Ala His Ile Gln Gly Asp Gly Val
 420 425 430
 Gly Ala Ile Cys Leu Gln Asp Thr Leu Ser Tyr Thr Pro Gln Ser Gly
 435 440 445
 Asn Val Asn Leu Ile Ile Gln Asn Leu Asp Thr Gln Val Leu Leu Ala
 450 455 460
 Ala Leu Pro Ser Asp Ile His Trp Lys Ser Met Leu Asn Gly Arg Ile
 465 470 475 480
 Lys Ala Met Trp Gln Ala Gly Gln Ser Pro Leu Val Asp Ala Val Leu
 485 490 495
 Tyr Ser Asp Asp Gly Thr Ile Gly Leu Thr Gln Glu Glu Thr Gly Tyr
 500 505 510
 Ile Glu Met Pro Tyr Gln Arg Ala Ser Val Ile Ala Lys Ser Val Asp
 515 520 525
 Asn Gly Leu Lys Val Arg Thr Asp Ile Leu Gly Thr Ala Gly Arg Gly
 530 535 540
 Tyr Ala Asp Val Ile Ile Asn Pro Lys Gln Thr Asp Lys Pro Ile Ser
 545 550 555 560
 Gly Ala Leu Val Met Asn Asp Leu Asn Leu Ala Val Leu Arg Pro Phe
 565 570 575
 Phe Pro Ser Ile Gln Thr Leu Ser Gly Lys Val Ser Leu Ala Gly Gly
 580 585 590
 Leu Gly Gly Thr Leu Ser Lys Pro Leu Phe Tyr Gly Asn Ala Asn Leu
 595 600 605
 Lys Asn Gly Ala Leu Ala Ile Val Gly Val Pro Met Pro Ile Ser Asp
 610 615 620


```

Val Asp Ala Thr Val Asn Ile Arg Gly Thr Lys Ala Arg Leu Asp Gly
625          630          635          640
Arg Phe Met Gly Gly Glu Gly His Gly Val Leu Tyr Gly Glu Met Asp
          645          650          655
Trp Ala Glu Glu Leu Tyr Ala Arg Leu Gly Val Phe Gly Glu Asn Leu
          660          665          670
Thr Val Ser Gln Pro Pro Leu Val Thr Ala Gln Ile Ser Pro Glu Leu
          675          680          685
Glu Val Ile Ile Arg Pro Phe Gln Lys Phe Val Asp Val Gln Gly Val
          690          695          700
Val Ser Ile Pro Ser Ala Thr Ile Arg Pro Pro Glu Ala Thr Ala Asp
705          710          715          720
Ile Val Thr Glu Ser Pro Asp Val Ser Val Ile Asp Arg Arg Ile Thr
          725          730          735
Gly Asn Ile Asp Gln Ile Leu Ser Arg Ala Lys Pro Trp Asp Ile Asn
          740          745          750
Ala Asn Ile Gly Val Asp Leu Gly Asn Glu Ile Glu Phe Arg Gly Phe
          755          760          765
Gly Ala Val Leu Pro Leu Ala Gly Ala Ile His Leu Thr Gln Ser Gly
          770          775          780
Gln Gly Ala Met Gln Ala Leu Gly Val Val Gln Val Ser Lys Arg Thr
785          790          795          800
Lys Ile Asp Val Ile Gly Gln Asn Leu Asp Leu Asn Tyr Ala Gln Ile
          805          810          815
Arg Phe Asp Gly Asp Met Leu Asn Pro Arg Leu Ser Ile Glu Gly Glu
          820          825          830
Lys Gln Ile Glu Gly Gln Thr Val Gly Val Arg Ile Arg Gly Thr Ala
          835          840          845
Ser Asn Pro Asn Ile Thr Val Phe Asn Asp Ala Gly Leu Asp Glu Tyr
          850          855          860
Gln Ala Met Asn Ala Leu Val Thr Gly Arg Ile Ser Glu Ser Ser Asp
865          870          875          880
Leu Gly Ile Thr Glu Gln Gly Phe Arg Ser Gln Val Thr Asn His Leu
          885          890          895
Ala Ala Ala Gly Leu Ser Leu Gly Leu Ser Gly Thr Arg Asp Leu Thr
          900          905          910
Asn Gln Ile Gly His Ala Phe Gly Phe Gln Ser Leu Thr Ile Asp Ala
          915          920          925
Ser Gly Ser Ser Asp Asp Thr Asn Val Asn Val Thr Gly Tyr Ile Thr
          930          935          940
Pro Asp Leu Tyr Leu Arg Tyr Gly Val Gly Val Phe Asn Ala Glu Ser
945          950          955          960
Thr Leu Ser Met Arg Tyr Gln Leu Thr Arg Arg Val Tyr Ile Glu Ala
          965          970          975
Thr Ser Ala Ala Glu Asn Met Val Asp Val Ile Tyr Arg Trp Lys Phe
          980          985          990

```

```

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> primer

```

```

<400> 5
cccaccgttt gcccttcaat

```

<210> 6
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 6
catttttccc gagcattcaa ac 22

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 7
aatcagccac ttatcgccac 20

<210> 8
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 8
gtaaaacgac ggccagt 17

<210> 9
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 9
caggaaacag ctatgac 17

<210> 10
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 10
tcatgaatga ctgaggcaaa g 21

<210> 11
<211> 25
<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 11

agatctaaac ttccaacgat aaatc

25

<210> 12

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 12

aaacaaatcg caccacgcc

20

<210> 13

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 13

acaaattgca gcgcattgtt gg

22